UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

LUCIMARA CRISTINA NAKATA SCADUTO

Estudo de sensibilidade ao alinhamento e desenvolvimento de uma metodologia para alinhamento de sistemas ópticos por meio da análise de aberrações de frente de onda utilizando redes neurais artificiais

São Carlos
2013

LUCIMARA CRISTINA NAKATA SCADUTO

Estudo de sensibilidade ao alinhamento e desenvolvimento de uma metodologia para alinhamento de sistemas ópticos por meio da análise de aberrações de frente de onda utilizando redes neurais artificiais

Tese apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Doutor em Ciências.

Área de Concentração: Física Aplicada

Orientador: Prof. Dr. Jarbas Caiado de Castro Neto

Versão Corrigida

(Versão original disponível na Unidade que aloja o Programa)

São Carlos
2013

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pelo Serviço de Biblioteca e Informação do IFSC,
com os dados fornecidos pelo(a) autor(a)

Scaduto, Lucimara Cristina Nakata
Estudo de sensibilidade ao alinhamento e desenvolvimento de uma metodologia para alinhamento de sistemas ópticos por meio da análise de aberrações de frente de onda utilizando redes neurais artificiais / Lucimara Cristina Nakata Scaduto; orientador Jarbas Caiado de Castro Neto - versão corrigida -- São Carlos, 2013.
216 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2013.

1. Alinhamento. 2. Aberrações. 3. Polinômios de Zernike. 4. Redes neurais artificiais. 5. Sistema óptico. I. Castro Neto , Jarbas Caiado de, orient. II. Título.

Aos meus pais Yochie e Antonino, e ao meu noivo Alan, com amor e gratidão, pelo incentivo, dedicação, compreensão e companheirismo.

AGRADECIMENTOS

Ao orientador Prof. Dr. Jarbas Caiado de Castro Neto, pela orientação e oportunidade de desenvolver este trabalho.

Ao Dr. Mario Antonio Stefani, diretor de P&D da Opto Eletrônica S.A., pelo incentivo, auxílio técnico, confiança e pela oportunidade de desenvolver este trabalho.

À Opto Eletrônica S.A., pela disponibilização da infraestrutura necessária para o desenvolvimento deste trabalho e por ceder os *software* MATLAB® e ZEMAX®.

À fundação CAPES pelo auxílio financeiro durante estágio de doutorado sanduíche, processo BEX 1098/11-0, através do Programa de Doutorado no Brasil com Estágio no Exterior (PDEE).

Ao Prof. Dr. José Sasián pela orientação durante o estágio de doutorado sanduíche e pelas valiosas discussões, que foram essenciais para o meu aprendizado e para o desenvolvimento deste trabalho.

Ao Prof. Dr. Zhao Liang pela ajuda e discussões que muito contribuiram para este trabalho.

A Wilson Mendes, José Otoboni, Vitor Montes, José Augusto Stuchi, Tatiana Pazelli, Fátima Yasuoka e Bráulio Albuquerque por toda ajuda e pelas valiosas discussões.

À Secretaria de Pós-Graduação, assessorada por Ricardo, Patrícia e Silvio, por toda a ajuda e atenção.

Ao Alan, pela compreensão e paciência ao longo de todos estes anos e pelo incentivo e motivação nos momentos mais difíceis.

Aos meus pais pelo apoio irrestrito em todos os momentos, pelo incentivo e compreensão.

Ao meu irmão Rogério e à Valdaisa pelo apoio e ajuda.

Aos meus familiares e amigos próximos pelo apoio e incentivo.

Aos meus sobrinhos Gabriel e Yasmin, pela alegria e descontração.

À Nina, pela alegria e companheirismo incondicional.

A todos que, de alguma forma, contribuiram para a realização deste trabalho.

RESUMO

SCADUTO, L. C. N. **Estudo de sensibilidade ao alinhamento e desenvolvimento de uma metodologia para alinhamento de sistemas ópticos por meio da análise de aberrações de frente de onda utilizando redes neurais artificiais.** 2013. 216 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Erros de alinhamento em sistemas ópticos não criam novas aberrações, mas alteram a dependência com o campo das aberrações já conhecidas. Neste trabalho, a sensibilidade teórica ao alinhamento, de sistemas ópticos reflexivos compostos por dois elementos, foi avaliada em função das constantes cônicas dos espelhos. Dentre as diferentes configurações consideradas nesta análise, uma específica apresenta menor sensibilidade à descentralização do espelho secundário. A utilização da teoria de aberração de onda aplicável a sistemas plano-simétricos revelou que a escolha apropriada da constante cônica do espelho secundário faz com que coma uniforme de terceira ordem seja compensado quando esse elemento encontra-se descentralizado, fazendo com que esse sistema seja livre da aberração mais importante causada a ele por desalinhamentos, tornando-o menos sensível. Este trabalho apresenta uma metodologia de alinhamento baseada na análise da frente de onda transmitida por sistemas ópticos, que utiliza redes neurais artificiais para a estimativa dos erros de alinhamento. A frente de onda transmitida por um sistema óptico carrega informações das aberrações desse sistema, que podem ser descritas em termos dos polinômios de Zernike. Esses polinômios podem ser usados para a análise dos efeitos de erros de alinhamento nas aberrações do sistema. Redes neurais artificiais são empregadas na análise dos coeficientes dos polinômios de Zernike visando avaliar o tipo de desalinhamento e a sua magnitude. As estimativas teóricas dos desalinhamentos tanto em sistemas reflexivos como em sistemas refrativos são satisfatórias quando o sistema é considerado perfeito, ou seja, as superfícies ópticas de seus elementos não apresentam erros de forma e não há ruído nos dados avaliados. Na presença de defeitos de fabricação ocorre degradação no desempenho do estimador. Além de descentralização e inclinação, redes neurais artificiais são capazes de fornecer uma estimativa de erros de posicionamento axial dos elementos do sistema. Com base nos estudos realizados, acredita-se que redes neurais artificiais constituem uma alternativa promissora no alinhamento de sistemas ópticos complexos.

Palavras-chave: Alinhamento. Aberrações. Polinômios de Zernike. Redes neurais artificiais. Sistema óptico.

ABSTRACT

SCADUTO, L. C. N. **Alignment sensitivity analysis and development of an optical systems alignment methodology based on the analysis of wave aberrations utilizing artificial neural networks.** 2013. 216 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Although misalignments in optical systems do not generate new aberration forms, they change the field-dependence of the known ones. In this research, the sensitivity of two-mirror optical systems due to misalignments is evaluated in function of the conic constants of the mirrors. Among the different configurations considered in this study, a specific one has shown low sensitivity due to decenter misalignments. The application of the wave aberration theory for plane-symmetric optical systems has revealed that the proper choice of the secondary mirror conic constant allows third-order uniform coma to be compensated, leading to a less sensitive system, free from the most important misalignment-induced aberration. This thesis also presents an alignment methodology based on the analysis of the transmitted wavefront utilizing artificial neural networks to estimate alignment errors in the components of the system. The transmitted wavefront carries information about the aberrations in the optical system, which can be described in terms of Zernike polynomials. Such polynomials are used for the analysis of the effects of misalignments on the aberrations of the system. Artificial neural networks are employed in the analysis of the coefficients of Zernike polynomials and used to evaluate both type and magnitude of the misalignments. Theoretical misalignments estimated in reflexive and refractive optical systems are satisfactory for perfect systems, i.e., systems with no surface errors, and noiseless data. When surface imperfections are considered, the performance of the estimator is reduced. Besides decenter and tilt misalignments, artificial neural networks can estimate axial positioning errors of the elements in the system, therefore they are believed to be a promising alternative for the alignment of complex optical systems.

Keywords: Alignment. Aberrations. Zernike polynomials. Artificial neural network. Optical system.

LISTA DE FIGURAS

Figura 1 – Erros de alinhamento nos elementos ópticos; a) Sistema alinhado; b) Elemento descentralizado; c) Elemento inclinado; d) Elemento deslocado axialmente. 44

Figura 2 – Aberrações de Seidel. A) Lente simples com aberração esférica onde raios que atravessam a lente em regiões mais afastadas do eixo óptico formam imagem em pontos mais próximos da lente; B) Lente simples com coma: 1) Raios que passam por regiões mais afastadas do eixo óptico formam imagem em alturas diferentes no plano imagem dos raios que passam pelo eixo; 2) Imagem de um objeto pontual formada por um sistema que apresenta coma. C) Comparação de imagens de sistemas com Distorção: 1) Distorção *pincushion* ou positiva; 2) Distorção *barrel* (barril) ou negativa. D) Astigmatismo e curvatura de campo: 1) Sistema com astigmatismo; 2) Ilustração das superfícies imagem sagital e tangencial posicionadas em relação à superfície de Petzval para um sistema com astigmatismo Fonte: Adaptada de SMITH (17). 52

Figura 3 – Representação do vetor de campo efetivo. Fonte: THOMPSON (2). 63

Figura 4 – Pontos onde a aberração é zero: a) Coma; b) Astigmatismo. Fonte: THOMPSON (2). 65

Figura 5 – Polinômios de Zernike. Fonte: Adaptada de WYANT (49). 74

Figura 6 – Gráficos de densidade dos polinômios de Zernike. Fonte: Adaptada de WYANT (50). 75

Figura 7 – Superfície fabricada por método de polimento determinístico analisada em três diferentes regiões de frequências espaciais (arbitrariamente escolhidas pelos autores). Fonte: AIKENS, DEGROOTE e YOUNGWORTH (52). 78

Figura 8 – Domínio dos diferentes tipos de erros superficiais em relação à frequência espacial. Na figura, PSD significa *Power Spectral Density function* ou função de densidade espectral. Fonte:

Adaptada de KOTHA e HARVEY (54) e HARVEY, LEWOTSKY, KOTHA (57). 79
Figura 9 – Interferogramas das aberrações de primeira ordem e de combinações delas. Fonte: Adaptada de ZYGO Corp (60). 82
Figura 10 – Esquema ilustrativo de um interferômetro de Fizeau. Fonte: Adaptada de GEARY (61). 83
Figura 11 – Modelo de neurônio não linear. Fonte: HAYKIN (15). 85
Figura 12 – Representação gráfica das funções de ativação: Heaviside, linear, logística e tangente hiperbólica. 88
Figura 13 – Ilustração da rede Perceptron. Fonte: SILVA, SPATTI e FLAUZINO (64). 91
Figura 14 – Ilustração da rede Adaline. Fonte: SILVA, SPATTI e FLAUZINO (64). 93
Figura 15 – Interpretação geométrica da regra Delta. Fonte: SILVA, SPATTI e FLAUZINO (64). 96
Figura 16 – Diferença entre as soluções das redes Perceptron e Adaline. Fonte: SILVA, SPATTI e FLAUZINO (64). 96
Figura 17 – Ilustração da rede Perceptron multicamadas. Fonte: SILVA, SPATTI e FLAUZINO (64). 97
Figura 18 – Notação para algoritmo *backpropagation.* Fonte: SILVA, SPATTI e FLAUZINO (64). 99
Figura 19 – Telescópios com dois espelhos: a) Gregorian e b) Cassegrain. Fonte: Adaptada de SMITH (17). 102
Figura 20 – Esquema ilustrativo das nomenclaturas adotadas Fonte: SCADUTO et al. (67). 102
Figura 21 – O gráfico à esquerda apresenta a sensibilidade a perturbações do tipo descentralização para diferentes configurações com mesmos f, d e B. O gráfico à direita exibe a sensibilidade a perturbações do tipo inclinação para as mesmas configurações. A legenda apresenta o valor da constante cônica do espelho primário de cada configuração analisada. Fonte: Adaptado de SCADUTO et al. (67). .. 107
Figura 22 – Gráficos de coma total (a menos de um fator de escala) em todo o campo (waves, λ=0,55 μm) dado pela equação (78). A figura à esquerda apresenta curvas do coma em todo o campo para a

configuração *classical Cassegrain*, no estado alinhado, com descentralização [mm] e com inclinação [°] no espelho secundário, respectivamente, de cima para baixo. Os mesmos gráficos são apresentados na coluna central para a configuração RC e na coluna à direita para a configuração com menor sensibilidade (LS). ..109

Figura 23 – Gráficos de astigmatismo total (a menos de um fator de escala) em todo o campo (waves, λ=0,55 μm) dado pela equação (79). A figura à esquerda apresenta curvas de astigmatismo em todo o campo para a configuração *classical Cassegrain*, no estado alinhado, com descentralização [mm] e com inclinação [°] no espelho secundário, respectivamente, de cima para baixo. Os mesmos gráficos são apresentados na coluna central para a configuração RC e, na coluna à direita, para a configuração com menor sensibilidade (LS). 110

Figura 24 - Gráficos dos coeficientes dos polinômios de Zernike: A) Z_5 ; B) Z_6 ; C) Z_7; D) Z_8 e E) RMS (waves, λ=0.55 μm) em função da inclinação do espelho secundário para as configurações *classical Cassegrain*, RC e LS. Fonte: SCADUTO et al. (67). 111

Figura 25 – Gráficos dos coeficientes dos polinômios de Zernike: A) Z_5 ; B) Z_6 ; C) Z_7; D) Z_8 e E) RMS (waves, λ=0.55 μm) em função de descentralização do espelho secundário para as configurações *classical Cassegrain*, RC e LS. Fonte: SCADUTO et al. (67). 112

Figura 26 – Descentralização em uma superfície esférica. Fonte: SCADUTO et al. (67). 115

Figura 27 – Gráfico dos valores de k_S calculados pela equação (103) versus valores encontrados nas análises de perturbação para sistemas com diferentes combinações de *f*, *B* e *d*. 120

Figura 28 – Gráfico da posição do ponto livre de coma Z_{CFP} em função da constante cônica do espelho secundário em sistemas de dois espelhos. Fonte: SCADUTO et al. (67).....120

Figura 29 – Fluxograma do procedimento para alinhamento de sistemas ópticos por meio da análise dos polinômios de Zernike. 127

Figura 30 – Fluxograma do procedimento para alinhamento de sistemas ópticos por meio da análise dos polinômios de Zernike – análises teóricas. 128
Figura 31 – Diagrama para a extração dos dados de treinamento a partir do modelo teórico do sistema. 130
Figura 32 – Fluxograma para a extração dos dados de treinamento da RNA. 132
Figura 33 – Fluxograma do processo de treinamento das RNAs. 140
Figura 34 – Montagens experimentais possíveis para os sistemas analisados nesta tese, que formam imagens de objetos localizados no infinito; a) O sistema é posicionado de forma que um feixe colimado incida sobre ele e uma esfera de referência, côncava ou convexa, é usada para refletir o feixe de volta ao interferômetro; b) O sistema é posicionado de forma que o seu ponto de foco coincida com o ponto de foco do feixe do interferômetro; o feixe emergente é então colimado e um espelho plano de referência é usado para retrorrefletir o feixe. Fonte: ZYGO CORPORATION (74). 141
Figura 35 – Modelo de um telescópio do tipo *classical Cassegrain*. 145
Figura 36 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia I, treinamento 1. ... 149
Figura 37 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia I, treinamento 1. 150
Figura 38 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia I, treinamento 1. 151
Figura 39 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia II, treinamento 1. .. 152
Figura 40 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia II, treinamento 1. 152

Figura 41 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia II, treinamento 1. 153
Figura 42 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia III, treinamento 1. .. 154
Figura 43 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia III, treinamento 1. 155
Figura 44 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia III, treinamento 1. 155
Figura 45 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia IV, treinamento 1. . 156
Figura 46 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia IV, treinamento 1. 157
Figura 47 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia IV, treinamento 1. 158
Figura 48 – a) Frente de onda da superfície de uma lente em que *piston* e inclinação (*tilt*) foram descontados da imagem; b) Frente de onda da mesma superfície em que apenas *piston* foi descontado. Fonte: Opto Eletrônica S.A. 162
Figura 49 – Frentes de onda de superfícies de diferentes lentes cujas especificações de fabricação para erros superficiais eram de 0,25 λ de P-V. Fonte: Opto Eletrônica S.A. 163
Figura 50 – Exemplos de erros de forma gerados nas superfícies dos espelhos (a) primário e (b) secundário considerando coeficientes de Zernike (*Standard*) de 5 a 22 e RMS=1,5E-5 mm. 164
Figura 51 – Exemplos de erros de forma gerados nas superfícies dos espelhos (a) primário e (b) secundário considerando coeficientes de Zernike (*Standard*) de 5 a 231 e RMS=1,5E-5 mm. 165

Figura 52 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras do sistema perfeito referentes à estratégia III, para amostras vindas de cinco sistemas imperfeitos com deformações contendo coeficientes de Zernike de 5 a 22. 167
Figura 53 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras do sistema perfeito referentes à estratégia III, para amostras vindas de sistemas imperfeitos com deformações contendo coeficientes de Zernike de 5 a 231. 168
Figura 54 – Saídas da rede treinada apenas com dados provenientes do sistema perfeito diante da apresentação das amostras extraídas de sistemas imperfeitos. Casos de descentralização em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada. 169
Figura 55 – Saídas da rede treinada apenas com dados provenientes do sistema perfeito diante da apresentação das amostras extraídas de sistemas imperfeitos. Casos de erros de alinhamento do tipo inclinação em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada..... 170
Figura 56 – Coeficiente Z_5 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo, para o sistema original, e dois sistemas com deformações superficiais tratados conforme descrito anteriormente. 171
Figura 57 – Coeficiente Z_6 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo, para o sistema original, e dois sistemas com deformações superficiais..... 172
Figura 58 – Coeficiente Z_7 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo para o sistema original, e dois sistemas com deformações superficiais..... 172

Figura 59 – Coeficiente Z_8 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campos para o sistema original, e dois sistemas com deformações superficiais. 172
Figura 60 – Saídas da rede treinada com dados provenientes de sistemas com deformações superficiais diante da apresentação das amostras extraídas destes mesmos sistemas imperfeitos que não haviam sido apresentados durante treinamento. Casos descentralização em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada...174
Figura 61 – Esquema ilustrativo das entradas e saídas da RNA estimadora e da RNA corretora.....175
Figura 62 – Histograma de erros e gráfico de erro em função do número da amostra para as saídas da RNA_E, treinada com dados extraídos de um sistema perfeito, diante de amostras de cinco sistemas com deformações superficiais.....175
Figura 63 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede corretora treinada. 176
Figura 64 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede após passar pela rede corretora, para amostras vindas dos 45 sistemas imperfeitos cujas amostras foram apresentadas à RNA durante treinamento, porém estes são dados de teste que correspondem aos 15% dos dados que não foram apresentados durante treinamento da RNA corretora. 177
Figura 65 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede após passar pela rede corretora, para amostras vindas dos 5 sistemas imperfeitos e que não foram considerados durante treinamento da RNA corretora. 177
Figura 66 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada, treinamento conjunto 1. 181

Figura 67 – Histograma de erros e gráfico de erro em função do número da amostra para as cinco saídas da rede treinada com as amostras referentes à estratégia III – treinamento 1, dados de teste. 182
Figura 68 – Histogramas de erros para cada uma das cinco saídas para a RNA treinada com as amostras referentes à estratégia III, treinamento conjunto 1, dados de teste. 182
Figura 69 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada, treinamento conjunto 1. 183
Figura 70 – Histograma de erros e gráfico de erro em função do número da amostra para as cinco saídas da rede treinada considerando coeficiente relativo à aberração esférica, treinamento conjunto 1 – amostras de teste. 184
Figura 71 – Histogramas de erros para cada uma das cinco saídas para a RNA, treinamento conjunto 1. 185
Figura 72 – Sistema óptico refrativo projetado para sensoriamento remoto orbital. 186
Figura 73 – Gráfico de MTF do sistema em toda faixa espectral. 187
Figura 74 – Efeitos de desalinhamentos nas aberração do sistema; a) astigmatismo em todo o campo e b) coma em todo o campo, para o sistema alinhado e com desalinhamento do tipo descentralização nas lentes 3, 5 e 7, respectivamente, de cima para baixo; c) astigmatismo em todo o campo e d) coma em todo o campo, para o estado alinhado e para desalinhamentos do tipo inclinação nas lentes 3, 5 e 7. 190
Figura 75- Interferogramas gerados no Zemax® para o sistema alinhado e para desalinhamentos do tipo descentralização e inclinação nas lentes 3, 5 e 7, respectivamente. 191
Figura 76 – EQM em função das épocas de treinamento, treinamento 1. 192
Figura 77 – Histogramas de erros para os desalinhamentos nas lentes a) 3, b) 5 e c) 7, e para d) deslocamentos axiais. 193

LISTA DE TABELAS

Tabela 1 – Termos de segunda ordem da função de aberração de onda. 49

Tabela 2 – Termos de quarta ordem da função de aberração de onda. 49

Tabela 3 – Grupos de aberrações em sistemas plano-simétricos, de acordo com a definição de Sasián (3). 57

Tabela 4 – Polinômios circulares de Zernike ($x = \rho\cos\theta, y = \rho\text{sen}\theta, 0 \le \rho \le 1, 0 \le \theta < 2\pi$). 73

Tabela 5 – Polinômios Fringe de Zernike. 74

Tabela 6 – Combinação de constantes cônicas dos espelhos. 106

Tabela 7 – Constantes cônicas, termos asféricos do espelho primário, coma linear residual e astigmatismo quadrático para as configurações analisadas. 108

Tabela 8 – Polinômios de Zernike (FRINGE) que representam aberrações primárias que podem ser usados como dados de entrada para a RNA. 134

Tabela 9 – Descrição das saídas da RNA, para cada elemento. 136

Tabela 10 – Parâmetros do sistema óptico. 145

Tabela 11 – Descrição dos intervalos de desalinhamentos e amostragens adotadas. 146

Tabela 12 – Estratégias adotadas para o conjunto de dados de entrada para treinamento da rede. 147

Tabela 13 – EQM para cada uma das quatro saídas da RNA e valor total, média dos valores de cinco treinamentos e desvio dos valores de EQM total em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia I – amostras de teste. 149

Tabela 14 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das quatro saídas da RNA e valor total, treinamentos com as amostras referentes à estratégia I – amostras de teste. 150

Tabela 15 – EQM para cada uma das quatro saídas da RNA e EQM total, valores médios e desvio do EQM total em cada treinamento em

relação ao seu valor médio, treinamentos com as amostras referentes à estratégia II – amostras de teste. 152
Tabela 16 – EQM para cada uma das quatro saídas da RNA e valor total, valor médio e desvio do EQM total em cada treinamento em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia III – amostras de teste. 154
Tabela 17 – EQM para cada uma das quatro saídas da RNA e valor total, valores médios e desvio dos valores de EQM total, em cada treinamento, em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia IV – amostras de teste. 156
Tabela 18 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das quatro saídas da RNA e valor total, treinamentos com as amostras referentes à estratégia IV – amostras de teste. 157
Tabela 19 – Descrição dos intervalos de desalinhamento adotados na análise dos efeitos de deformações superficiais no desempenho da rede. 166
Tabela 20 – EQM para cada uma das quatro saídas da RNA diante de amostras de teste do sistema sem deformações e de sistemas com imperfeições geradas com coeficientes de Zernike nos intervalos de 5 a 22 e de 5 a 231. 168
Tabela 21 – EQM para cada uma das quatro saídas da RNA quando são apresentadas amostras provenientes de sistema com elementos contendo erros de superfície já conhecidos e desconhecidos pela rede. 177
Tabela 22 – Descrição dos intervalos de desalinhamentos e amostragens adotadas na análise incluindo deslocamentos axiais. 179
Tabela 23 – EQM para cada uma das cinco saídas da RNA e valor total, media dos valores de cinco treinamentos e desvio dos valores de EQM total em relação ao seu valor médio, amostras referentes a estratégia III– amostras de teste. 181
Tabela 24 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das cinco saídas da RNA, valor total e média, estratégia III – amostras de teste. 181
Tabela 25 – EQM para cada uma das cinco saídas da RNA e valor total, média dos valores de cinco treinamentos e desvio dos valores de EQM total

em relação ao seu valor médio, amostras referentes aos treinamentos considerando Z5, Z6, Z7, Z8 e Z9 medidos em cinco pontos no campo– amostras de teste. 184
Tabela 26 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das cinco saídas da RNA, valor total e média, amostras referentes aos treinamentos considerando Z5, Z6, Z7, Z8 e Z9 medidos em cinco pontos no campo– amostras de teste 184
Tabela 27 – Características do sistema óptico. 186
Tabela 28 – Tolerâncias de montagem e alinhamento do sistema – análise simplificada. 188
Tabela 29 – EQM em cada treinamento, valor médio em cinco treinamentos e desvio dos valores de EQM em relação ao seu valor médio. Porcentagem das amostras com erro, em módulo, inferior a 0,0025 mm ou 0,0025 °, valor médio e desvio do valor em cada treinamento em relação ao valor médio – amostras de teste. 193
Tabela 30 – Coeficientes das aberrações de frente de onda, para elementos inclinados, contribuição de uma superfície esférica e da capa asférica (3). 211

LISTA DE ABREVIATURAS E SIGLAS

AX	Deslocamento Axial;
BFL	do inglês, *Back Focal Length;*
CCD	do inglês, *Charge Coupled Device;*
CFP	do inglês, *Coma-Free Point;*
DDE	do inglês, *Dynamic Data Exchange*;
DWS	do inglês, *Differential Wavefront Sampling;*
DX	Descentralização no eixo *x*;
DY	Descentralização no eixo *y*;
EFL	do inglês, *Effective Focal Length;*
EQM	Erro Quadrático Médio;
FN	Falso Negativo;
FNR	do inglês, *False Negative Rate;*
FP	Falso Positivo;
FPR	do inglês, *False Positive Rate;*
LA	do inglês, *Local Axis;*
LS	do inglês, *Less Sensitive;*
MC	Matriz de Confusão;
MFWS	do inglês, *Multi-Field Wavefront Sampling;*
MLP	do inglês, *Multilayer Perceptron;*
MSF	do inglês, *Mid-Spatial Frequency Errors;*
NAT	do inglês, *Nodal Aberration Theory*;
OAR	do inglês, *Optical Axis Ray;*
OPD	do inglês, *Optical Path Difference;*
OSC	do inglês, *Offense Against the Abbe Sine Condition;*

PCA do inglês*, Principal Component Analysis*

PMC Perceptron Multicamadas;

PSD do inglês, *Power Spectral Density Function;*

PSF do inglês, *Point Spread Function;*

PV do inglês, *Peak to Valey;*

RA do inglês, *Reference Axis;*

RBF Redes de Base Radial;

RC Ritchey-Chrétien;

RMS do inglês, *Root Mean Square;*

RNA Rede Neural Artificial;

SE Saída Esperada;

SO Saída Obtida;

SVD do inglês, *Singular Value Decomposition;*

TCC Total de Classificações Corretas;

TCI Total de Classificações Incorretas;

TG Total de Genuínos;

TI Total de Impostores;

TMA do inglês, *Three-Mirror Anastigmatic;*

TNR do inglês, *True Negative Rate;*

TPR do inglês, *True Positive Rate;*

TS Total de Saídas;

TSN Total de Saídas Negativas;

TSP Total de Saídas Positivas;

TX Inclinação eixo *x*;

TY Inclinação eixo *y*;

SUMÁRIO

1 INTRODUÇÃO 31

1.1 Justificativa 31

1.2 Motivação 34

1.3 Objetivos 35

1.4 Erros de alinhamento em sistemas ópticos 36

1.5 Métodos para alinhamento de sistemas ópticos – breve histórico de trabalhos relacionados 37

1.6 Estrutura do Trabalho 41

2 REVISÃO BIBLIOGRÁFICA 43

2.1 Introdução 43

2.2 Aberrações em sistemas ópticos 45

2.2.1 Aberrações em sistemas ópticos com simetria rotacional 48

2.2.2 Aberrações de Seidel 50

2.2.3 Sistemas ópticos não rotacionalmente simétricos ou desalinhados 53

2.3 Polinômios de Zernike 66

2.4 Defeitos nas superfícies de elementos ópticos 76

2.4.1 Erros de forma 79

2.5 Métodos Interferométricos 81

2.6 Redes Neurais Artificiais - RNA 83

2.6.1 Introdução e conceitos básicos 83

2.6.2 Perceptron 90

2.6.3 *Toolbox* de redes neurais artificiais do Matlab 100

3 TELESCÓPIOS REFLEXIVOS COM DOIS ELEMENTOS – ANÁLISE DE SENSIBILIDADE AO ALINHAMENTO 101

3.1 Introdução 101

3.2 Telescópios reflexivos com dois espelhos – descrição de algumas configurações 101
3.3 Alinhamento de sistemas reflexivos com dois elementos 104
3.4 Sensibilidade ao alinhamento 105
3.5 A configuração de menor sensibilidade - LS 111
3.5.1 Contribuição esférica 113
3.5.2 Contribuição asférica 116
3.5.3 Termo total da aberração coma uniforme para sistemas reflexivos com dois elementos 117
3.6 O ponto livre de coma 119
3.7 Conclusões deste capítulo 121
4 METODOLOGIA PARA ALINHAMENTO DE SISTEMAS ÓPTICOS 123
4.1 Introdução 123
4.2 Descrição do problema 123
4.3 Proposta de metodologia para alinhamento de sistemas ópticos 124
4.4 Modelagem do sistema óptico 129
4.5 Extração das amostras de treinamento 130
4.5.1 Perturbação dos elementos do sistema e extração das amostras de treinamento 131
4.5.2 Entradas 133
4.5.3 Saídas 135
4.5.4 Amostras de validação e teste 136
4.5.5 Tratamento dos dados 137
4.6 Elaboração da RNA 138
4.7 Treinamento e avaliação do desempenho da RNA 139
4.7.1 Treinamento da RNA 139
4.7.2 Avaliação do desempenho da RNA 140
4.8 Montagem experimental e extração de dados reais 141

4.9 Processo iterativo para alinhamento de sistemas ópticos 142
5 RESULTADOS E DISCUSSÕES 143
5.1 Introdução 143
5.2 Sistema reflexivo com perturbações no espelho secundário 143
5.2.1 Modelagem do sistema óptico 144
5.2.2 Efeitos da escolha do conjunto de dados de entrada 146
5.2.3 Resultados e discussões 148
5.3 Efeitos de erros de forma nas superfícies dos elementos 160
5.3.1 Modelagem óptica das superfícies com erros de forma 160
5.3.2 Resultados e discussões 166
5.4 Sistema reflexivo com perturbações no espelho secundário e erros de posicionamento axial 178
5.4.1 Desalinhamentos no sistema 179
5.4.2 Resultados e discussões 180
5.5 Sistema refrativo com vários elementos 186
5.5.1 Descrição do sistema 186
5.5.2 Estimativa do estado de alinhamento 190
5.5.3 Resultados, discussões e conclusões 192
6 CONCLUSÕES GERAIS E SUGESTÕES PARA TRABALHOS FUTUROS ... 195
6.1 Conclusões 195
6.2 Sugestões para trabalhos futuros 197
REFERÊNCIAS 199
APÊNDICE A – Revisão da teoria de aberrações de onda aplicáveis a sistemas plano-simétricos 207
APÊNDICE B – Algoritmo *backpropagation* – descrição do processo de ajuste dos pesos sinápticos 213

1 INTRODUÇÃO

1.1 Justificativa

O desempenho final de um sistema óptico depende de diversos fatores, dentre eles, das aberrações presentes nesse sistema. Durante o processo de desenvolvimento, busca-se minimizar essas aberrações, para que o sistema final possa ter o desempenho especificado.

Erros ocasionados pelos processos de fabricação e montagem também influenciam o desempenho do sistema e devem ser considerados, desde a fase de projeto, de modo que a sensibilidade a esses erros seja conhecida e até mesmo minimizada, se possível.

As tolerâncias de fabricação e montagem devem ser especificadas para garantir o desempenho final do sistema. Requisitos relativos aos parâmetros construtivos asseguram que a imagem formada conterá aberrações que resultam em degradação de desempenho dentro de um limite aceitável. Além disso, devem ser avaliadas e especificadas as irregularidades superficiais que são aceitáveis para o sistema no que diz respeito às aberrações e ao espalhamento de luz (1).

O processo de alinhamento de um sistema óptico visa ajustar o posicionamento de seus elementos, de modo a minimizar as aberrações do sistema devidas a erros de alinhamento, até que este atinja o desempenho esperado.

É importante saber o estado de alinhamento do sistema. Uma forma de avaliar isso é através da análise das aberrações de frente de onda do sistema. Essas aberrações são modificadas pela presença de erros de alinhamento em seus elementos (2-3).

O uso de interferometria no processo de alinhamento de sistemas ópticos vem sendo proposto na literatura. A seção 1.5 apresenta um breve histórico de trabalhos relacionados.

Interferogramas possibilitam a observação das aberrações presentes em um sistema óptico. Estas aberrações podem ser descritas em termos dos polinômios de Zernike (2, 4).

Diversas metodologias foram propostas para a estimativa dos erros de alinhamento presentes em um sistema a partir das aberrações. Trabalhos publicados por diversos autores indicam bons resultados em várias delas. Uma breve revisão desses trabalhos é apresentada na seção 1.5.

Em sistemas já bem conhecidos, como telescópios reflexivos com dois espelhos, ou mesmo em sistemas com três espelhos conhecidos como TMA (*Three mirror anastigmatic*), existem soluções analíticas capazes de descrever as aberrações causadas por erros de alinhamento (5-7).

Em sistemas mais complexos, com mais elementos, contendo muitos graus de liberdade, os efeitos das aberrações não são facilmente separáveis. Nesses casos, métodos numéricos vêm sendo aplicados, dentre eles o SVD (*Singular Value Decomposition*) (6).

O software CODE V possui uma ferramenta para alinhamento óptico chamada ALI, que utiliza SVD para a análise dos erros de alinhamento de sistemas ópticos (8). Essa ferramenta utiliza medidas da frente de onda do sistema fabricado, obtidas por meio de interferômetros, que são importadas e utilizadas nas análises que fornecem indicações dos ajustes necessários para a correção ou redução das aberrações oriundas de erros de alinhamento.

Neste trabalho, a sensibilidade de diferentes configurações de sistemas com dois espelhos foi analisada em função de erros de alinhamento do tipo inclinação e descentralização do espelho secundário. Foram estudados os efeitos que esses erros ocasionam nas aberrações desses sistemas e a influência da escolha das constantes cônicas na sensibilidade do sistema, com base na teoria de aberrações de frente de onda aplicável a sistemas plano-simétricos. Além disso, propõe-se o uso de redes neurais artificiais para a estimativa de erros de alinhamento presentes em sistemas ópticos por meio da análise das aberrações do sistema.

Redes neurais artificiais têm sido aplicadas nos mais variados tipos de problemas; no campo de óptica, pode-se citar, por exemplo, o uso de redes neurais para a estimativa de distorção de fase causada por turbulência atmosférica em imagens de

telescópios terrestres e para controle de óptica adaptativa em telescópios astronômicos (9-11). Além disso, foram aplicadas redes neurais artificiais para a determinação das aberrações do telescópio espacial Hubble por meio de imagens estelares (12).

Há relatos na literatura da utilização de redes neurais artificiais para alinhamento óptico de lasers e sistemas ópticos (13-14). Hart, Bailey e Palumbo (13) propõem o uso de redes neurais artificiais para alinhamento dos componentes ópticos de um laser químico de alta potência através de informações de posição, padrão e intensidade de pico da imagem. Yoon et al. (14) propõem o uso de redes neurais artificiais para o alinhamento de um sistema composto por dois elementos asféricos de tamanhos muito diferentes com tolerâncias críticas de alinhamento. Na metodologia proposta por eles, redes neurais são usadas para o reconhecimento de padrões interferométricos. Os dados de entrada usados foram as distribuições de intensidades em pequenas regiões em que foi dividido o padrão interferométrico, e o número de franjas ao longo de determinadas linhas no padrão.

Apesar de a literatura apresentar diferentes propostas de metodologias de alinhamento óptico que empregam os polinômios de Zernike, não foram encontrados relatos do uso de redes neurais artificiais em que as informações fornecidas para seu treinamento fossem baseadas no uso dos polinômios de Zernike.

A abordagem proposta neste trabalho se diferencia das demais já que utiliza, como dados de entrada, alguns coeficientes deste polinômio relacionados a aberrações presentes no sistema que sofrem influência dos erros de alinhamento. A escolha dos coeficientes usados tem como base a teoria de aberrações de onda.

A grande variedade de aplicações das redes neurais artificiais deve-se ao fato de possuirem a capacidade de armazenar conhecimento, adquirido a partir de seu ambiente através de um processo de treinamento, e torná-lo disponível para uso (15).

Análises teóricas indicam que a escolha apropriada dos coeficientes dos polinômios de Zernike que compõem os dados de entrada da rede possibilita a estimativa de erros de alinhamento do tipo descentralização, inclinação e deslocamento axial.

Os métodos para alinhamento óptico propostos na literatura, que utilizam os polinômios de Zernike, abordam apenas descentralização e inclinação.

Deslocamentos axiais podem ser estimados por meio do uso da ferramenta ALI, do software CODE V (16).

A proposta apresentada nesta tese, além de utilizar redes neurais artificiais na análise das aberrações do sistema induzidas por erros de alinhamento, por meio dos polinômios de Zernike, é também capaz de estimar deslocamentos axiais. Os efeitos causados por erros de forma, que ocorrem durante o processo de fabricação, no desempenho do estimador também foram abordados.

Técnicas de alinhamento óptico baseadas nas aberrações de frente de onda são de grande importância quando se trata de sistemas ópticos complexos. A aplicação dessas metodologias em sistemas reais podem proporcionar um alinhamento mais preciso e ocasionar minimização do tempo envolvido no processo de alinhamento.

1.2 Motivação

A principal motivação para o desenvolvimento desta metodologia de alinhamento óptico vem da experiência em projetos anteriores, realizados na empresa Opto Eletrônica S.A., onde foram desenvolvidos, fabricados e integrados sistemas ópticos de alta qualidade de imagem, empregados em sensoriamento remoto orbital.

Dentre os sistemas desenvolvidos, pode-se citar como exemplo a câmera MUX (*Multispectral Camera*) dos satélites CBERS 3 & 4. Devido aos requisitos de projeto, o sistema óptico desenvolvido apresentou tolerâncias críticas de fabricação e grande sensibilidade ao posicionamento relativo entre os componentes ópticos.

Apesar das câmeras apresentarem muitos recursos mecânicos desenvolvidos para a finalidade de alinhamento, o processo demandava considerável tempo para concretização, devido à lenta convergência do sistema.

Medidas interferométricas da frente de onda transmitida pelo sistema, realizadas ao longo do processo de alinhamento indicavam que havia elementos desalinhados, porém, em um sistema com 12 elementos, a tarefa de caracterização dos desvios é muito complexa.

A demanda de sistemas ópticos com maior resolução e alto desempenho, além de faixas espectrais extensas, e outros requisitos específicos de projeto, torna necessário o desenvolvimento de sistemas ópticos reflexivos com elementos asféricos fora do eixo. Neste caso, a dificuldade de alinhamento torna-se ainda maior.

É dentro deste contexto que o estudo das influências dos desvios de posicionamento relativo dos componentes em sistemas ópticos ganha grande importância para o desenvolvimento de novas metodologias de alinhamento.

O desenvolvimento destas novas metodologias de alinhamento viabiliza o desenvolvimento e a construção de sistemas ópticos mais complexos.

1.3 Objetivos

Este trabalho tem como objetivo analisar a sensibilidade teórica ao alinhamento de sistemas ópticos e entender os efeitos causados por erros de alinhamento nas aberrações ópticas. Além disso, este trabalho apresenta uma metodologia de alinhamento óptico baseada na análise das aberrações do sistema.

Essa metodologia baseia-se na aplicação de redes neurais artificiais para a análise dos coeficientes dos polinômios de Zernike, ajustados à frente de onda transmitida pelo sistema, para a estimativa dos erros de alinhamento presentes nele.

Os estudos teóricos e análises, focos desta tese, visam avaliar o desempenho de redes neurais artificiais quando empregadas na estimativa do estado de alinhamento de sistemas ópticos reflexivos e refrativos. Os efeitos causados por irregularidades superficiais no desempenho do estimador também são avaliados.

Como exemplo de aplicação, a metodologia aqui proposta foi teoricamente aplicada a um sistema reflexivo com dois elementos, e ao sistema óptico da câmera MUX para a estimativa dos desalinhamentos de 3 lentes dentre as 12 que o sistema possui. Aplicações práticas da metodologia proposta serão realizadas em trabalhos futuros.

1.4 Erros de alinhamento em sistemas ópticos

As aberrações presentes em sistemas ópticos axialmente simétricos são descritas pelas chamadas aberrações primárias ou aberrações de Seidel (5, 17). Quando algum elemento do sistema está inclinado ou descentralizado, ou seja, desalinhado, o sistema deixa de ter simetria axial e as aberrações de Seidel já não podem mais descrever corretamente as aberrações deste sistema. Diversos autores têm estudado as aberrações presentes em sistemas sem simetria rotacional ao longo dos anos (2-3, 18-19).

Em 1976, Richard A. Buchroeder (19) estudou sistemas ópticos com elementos inclinados e, como resultado de seu trabalho, chegou à conclusão de que as aberrações no plano imagem de sistemas sem simetria rotacional ou desalinhados continuam sendo a soma das contribuições individuais das superfícies; contudo, essas contribuições individuais não possuem mais um centro comum sobre o eixo óptico. Desenvolvendo as conclusões do trabalho de Buchroeder, Roland V. Shack e Kevin P. Thompson escreveram as funções de aberração na forma vetorial, utilizando para isso a multiplicação vetorial, incluindo as contribuições de componentes inclinados. Chegaram, assim, a um formalismo que descreve apropriadamente os efeito causados por elementos inclinados nas aberrações dos sistemas ópticos e estabeleceram o conceito de nodos e campos de aberrações.

Em sistemas desalinhados ou sem simetria axial, nenhuma aberração nova é criada, porém as aberrações já conhecidas ganham novas dependências com o campo. Estas novas dependências apresentam ordens menores que as já conhecidas, ou seja, uma aberração com dependência linear com o campo, como o coma, desenvolverá termos com dependência uniforme com o campo na presença de elementos desalinhados (20).

R. Buchroeder, R. Shack e K. Thompson explicaram que, quando os elementos de um sistema estão desalinhados, a contribuição de cada superfície dele para o campo de aberrações continua sendo rotacionalmente simétrica em torno de algum ponto no campo. No caso de superfícies asféricas, o deslocamento do centro de simetria da parte esférica da superfície e o da “capa” asférica serão diferentes

apesar de ambos manterem a sua natureza simétrica em torno do centro deslocado, mas este ponto não está mais localizado sobre o eixo óptico (20-23).

A teoria desenvolvida por R. V. Shack e K. Thompson, conhecida como NAT (*Nodal Aberration Theory* ou Teoria de Aberração de Onda Nodal), também prevê a localização no campo onde as aberrações são nulas (2).

Uma abordagem muito útil, unidimensional, aplicável a sistemas plano-simétricos foi desenvolvida por J. Sasián (3). Em seu trabalho, ele pôde também notar que, além de nodos, linhas de nodos podiam existir.

Dessa forma, estas teorias descrevem os efeitos dos erros de alinhamento nos campos de aberrações de sistemas ópticos e explicam o que ocorre com as aberrações quando diferentes tipos de erros de alinhamento estão presentes no sistema. No próximo capítulo, essas duas teorias serão apresentadas com detalhes.

1.5 Métodos para alinhamento de sistemas ópticos – breve histórico de trabalhos relacionados

Neste tópico é apresentado um breve resumo dos trabalhos prévios encontrados na literatura que utilizam a análise da frente de onda e métodos de alinhamento auxiliados por computador (*computer-aided alignment methods*) como ferramentas para o alinhamento de sistemas ópticos.

Os primeiros trabalhos que buscam desenvolver uma metodologia eficiente de alinhamento de sistemas ópticos utilizando métodos auxiliados por computador datam da década de 80, e as suas primeiras aplicações foram para alinhamento de sistemas reflexivos fora do eixo (24).

Um dos métodos propostos é conhecido como método da otimização reversa, que consiste em perturbar o modelo teórico de um sistema óptico ideal (perfeitamente alinhado), de forma a obter um sistema teórico com as mesmas características do sistema em teste imperfeito (*as built*). Se os dados que definem o desempenho do sistema correspondem a um único conjunto de elementos com desvios de

alinhamento específicos, então o sistema modelo terá os mesmos desalinhamentos que o sistema em teste que se deseja alinhar (25).

O nome dado a esse método, “otimização reversa”, origina-se do fato de, normalmente, serem usadas funções de mérito padrões de otimização de sistemas ópticos para a obtenção desses sistemas modelo. Essas funções são criadas, a princípio, para melhorar o sistema, não para torná-lo pior, como ocorre nesta aplicação (25).

Os dados que definem o desempenho do sistema podem ser obtidos de diferentes formas, tais como medidas da frente de onda ou das aberrações, ou qualquer outro parâmetro que forneça dados suficientes e possa ser tomado como “objetivo” durante o processo de otimização. As variáveis usadas no processo de otimização do sistema são parâmetros tais como inclinação, descentralização e posicionamento de todos ou de alguns elementos do sistema (25). Desta forma, a otimização reversa, por meio da variação de parâmetros específicos do sistema, como inclinação, descentralização e posicionamento de determinados elementos, possibilita a reprodução dos dados obtidos por meio das medidas do sistema real, levando, assim, à obtenção de um sistema teórico com o mesmo desempenho do real e permitindo que sejam descobertos quais elementos estão desalinhados dentro do sistema.

Lundgren e Wolfe (25) propuseram o uso desse método para o alinhamento de um telescópio reflexivo fora do eixo composto por três espelhos e obtiveram sucesso. Nesse trabalho, os autores obtiveram as aberrações de raios do sistema, e a otimização reversa foi feita utilizando-se o software ACCOS V® por meio do método dos mínimos quadrados. Otimização reversa para alinhamento de telescópio de três espelhos foi proposta previamente, em 1987, por Hwan J. Jeong, Geoge N. Lawrence e Kie B. Nahm (26).

Em um outro trabalho, Lundgren e Wolfe (27) propuseram o uso desta metodologia não apenas para a determinação dos erros de alinhamento de sistemas ópticos, mas também para a avaliação dos erros das superfícies ópticas dos elementos (*surface figures*) que apresentam alguma dificuldade para medidas interferométricas. Os autores demonstraram que o método de otimização reversa necessita de um modelo acurado do sistema, fato que pode ser usado para a detecção e compensação de

erros de fabricação desses elementos. Ainda neste trabalho, os autores apresentaram brevemente a aplicação desse método para um sistema de três lentes e um telescópio.

No trabalho descrito por Yiafan Huang, Lin Li e Yinhua Cao (28), os autores descrevem um método de alinhamento auxiliado por computador diferente do anterior. Nesse trabalho, os parâmetros de alinhamento do sistema são encontrados por meio do uso de uma matriz de sensibilidade do sistema, que relaciona os erros de alinhamento presentes no sistema com as aberrações de frente de onda. São apresentadas também algumas informações a respeito da obtenção da matriz de sensibilidade do sistema por meio da escolha de coeficientes dos polinômios de Zernike específicos para ele. Os autores aplicaram o método de decomposição em valores singulares para a redução e consequente otimização da matriz de sensibilidade.

Bin Zhang et al. (29) fizeram uso deste método, usando mínimos quadrados para o alinhamento de um sistema reflexivo com três elementos, asféricos, fora do eixo. Os autores utilizaram os erros de frente de onda obtidos por meio de inteferogramas em diferentes pontos ao longo do campo de visada. As aberrações foram obtidas das medidas por meio dos polinômios de Zernike.

Xiaofei Yang, ChangYuan Han, Jingchi Yu (24) também aplicaram o método ao alinhamento de um sistema óptico reflexivo fora do eixo composto por três espelhos. Nesse trabalho, os autores se basearam nas linearidades do sistema e aplicaram o método da matriz de sensibilidade e o método dos mínimos quadrados no processo de alinhamento auxiliado por computador, mostrando que os resultados finais obtidos com esse método eram superiores aos obtidos utilizando-se as ferramentas de alinhamento do software CODE V®.

Albuquerque et al. (30) apresentaram uma descrição detalhada de uma metodologia aplicando estimadores bayesianos para a determinação dos erros de alinhamento em sistemas refrativos complexos. Os polinômios de Zernike foram utilizados nesse trabalho, que incluiu também termos não lineares. Os autores aplicaram o método inicialmente a um sistema simples e apresentaram uma proposta de aplicação da técnica para o alinhamento do sistema óptico da câmera multiespectral MUX dos satélites sino-brasileiros CBERS 3&4.

Lee et al. (31) se propuseram a formalizar a influência do alinhamento na frente de onda de um sistema óptico em termos de fase e modulação de amplitude. Os autores estudaram como as diferentes combinações de erros de alinhamento interelementos influenciam nas aberrações do sistema e encontraram uma função que descreve estes acoplamentos. Observaram que os mesmos padrões de aberrações podem ser criados por diferentes elementos, porém com diferentes magnitudes, e que estes podem se somar ou cancelar uns aos outros. Esse fato indica que há muitas formas de compensar tais aberrações, mas também indica que o processo de alinhamento de um sistema de muitos elementos pode ser muito complicado devido à existência de mínimos locais nos métodos de otimização. A conclusão a que chegaram foi que as aberrações de ambos, campo e alinhamento, apresentam as mesmas características e que estas podem ser modeladas da mesma forma.

Uma nova metodologia foi proposta por Lee et al. (32). A metodologia proposta por eles difere das anteriores, basicamente, pela forma como os dados de entrada para a otimização do sistema, os “alvos”, são apresentados. Segundo os autores, trabalhos anteriores propunham o uso de amostragem de frente de onda em múltiplos pontos ao longo do campo de visada do sistema, método chamado MFWS (*Multi-Field Wavefront Sampling*). Tal método pode apresentar baixo desempenho de estimação de desalinhamentos em sistemas de muitos elementos devido, segundo os autores, a degenerescências intrínsecas e/ou devido à natureza não linear dos acoplamentos interelementos. Dessa forma, os autores propõem um método chamado de DWS (*Differential Wavefront Sampling*), que leva em conta os efeitos de alinhamento interelementos. Tal abordagem apresentou resultados superiores ao método MFWS, mostrando ser mais precisa e convergindo mais rapidamente, sendo confiável mesmo quando foram inseridos ruídos nas medidas e na presença de deformações superficiais simétricas de elementos, por exemplo, desvios do valor do raio de curvatura. Tal abordagem pode ser influenciada pela presença de aberrações astigmáticas e comáticas nas superfícies dos elementos, fato que pode confundir o estimador e reduzir sua eficiência.

Lee et al. (33) descreveram analiticamente os efeitos do alinhamento interelemento em sistemas ópticos com pupila circular e concluiram que o desalinhamento de uma superfície altera as características de aberração dela e das superfícies

subsequentes. Mostraram que descentralização e inclinação produzem deslocamento lateral das coordenadas de campo e pupila[a], e que deslocamentos axiais levam a modificações nos coeficientes de aberração alterando fator de escala e distâncias paraxiais.

Podem ser citados ainda diversos outros trabalhos que aplicam estas técnicas e/ou propõem pequenas variações, basicamente na forma em que os dados (objetivos) são obtidos do sistema que se busca alinhar (obtenção das funções de aberrações, interferometria, método de Hartmann), na forma como estes dados são apresentados ao estimador (coeficientes dos polinômios de Zernike, diferenças dos valores reais e teóricos, amostragem em múltiplos pontos ao longo do campo de visada, amostragem diferencial) etc. Algumas comparações entre as diferentes abordagens foram feitas (24, 32).

Polinômios de Zernike dependentes do campo foram propostos por Tessieres (34) para o alinhamento de sistemas ópticos. Manuel (6) dá continuidade a essa abordagem, propondo polinômios de Zernike dependentes do campo ortogonais.

Em sua tese, Manuel aplica o método proposto por Tessieres ao alinhamento de um telescópio fora do eixo com dois espelhos. Combinando características únicas desses sistemas, como os seus dois pontos neutros, com os coeficientes dos polinômios de Zernike com dependência com o campo, foi possível construir uma matriz de sensibilidade para o alinhamento do sistema.

1.6 Estrutura do Trabalho

Dados os objetivos deste trabalho, considera-se, além do capítulo introdutório, o desenvolvimento de mais cinco capítulos apresentando a seguinte estrutura:

Capítulo 2 – **Revisão bibliográfica** – apresenta uma introdução a respeito das aberrações em sistemas ópticos, polinômios de Zernike, métodos interferométricos e redes neurais artificiais.

[a] Pupila: Dada pela imagem do *STOP* de abertura quando visto do espaço objeto (pupila de entrada) e do espaço imagem (pupila de saída). Fonte: The Photonics Dictionary, Pittsfield: Laurin Publishing, p. D113, 2004.

Capítulo 3 – **Telescópios reflexivos com dois elementos – análise de sensibilidade ao alinhamento** – apresenta uma revisão sobre as configurações mais conhecidas de telescópios com dois espelhos, analisa e compara a sensibilidade ao alinhamento de diferentes configurações, analisa uma configuração particular que exibe menor sensibilidade a erros de alinhamento do tipo descentralização e justifica este comportamento com base na teoria de aberração de onda.

Capítulo 4 – **Metodologia para alinhamento de sistemas ópticos** – apresenta a metodologia proposta para alinhamento óptico por meio das aberrações de frente de onda.

Capítulo 5 – **Resultados e discussões** – apresenta os resultados da aplicação teórica do método proposto em sistemas ópticos com dois espelhos e em sistemas refrativos complexos. Apresenta uma análise dos efeitos de deformações superficiais nos resultados fornecidos pela RNA.

Capítulo 6 – **Conclusões gerais e sugestões para trabalhos futuros** – são apresentadas conclusões do trabalho e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

2.1 Introdução

Este capítulo tem como objetivo apresentar uma revisão dos temas considerados relevantes para este trabalho.

Os seguintes tópicos serão abordados:

- Aberrações em sistemas ópticos:
  - Sistemas com simetria rotacional;
  - Sistemas sem simetria rotacional ou desalinhados;
- Polinômios de Zernike;
- Erros de superfície em componentes ópticos;
- Métodos interferométricos;
- Redes neurais artificiais – RNA.

Neste trabalho, os erros de alinhamento abordados são definidos como:

- DESCENTRALIZAÇÃO (*Decenter*):
  Elementos descentralizados[b] (ou descentrados) possuem seus vértices deslocados do eixo óptico do sistema e o eixo óptico desse elemento está paralelo ao do sistema (Figura 1, b).
- INCLINAÇÃO (*Tilt*):
  O eixo óptico de um elemento inclinado não é mais paralelo ao eixo do sistema óptico. Esses eixos formam um ângulo de inclinação entre si (Figura 1, c).

  Definições mais detalhadas podem ser encontradas na literatura (35).

[b] De acordo com o Vocabulário Ortográfico da Língua Portuguesa, Academia Brasileira de Letras, 5° edição, 2009, pag. 260, coluna 4, pode-se escrever descentrado ou descentralizado, descentralização.

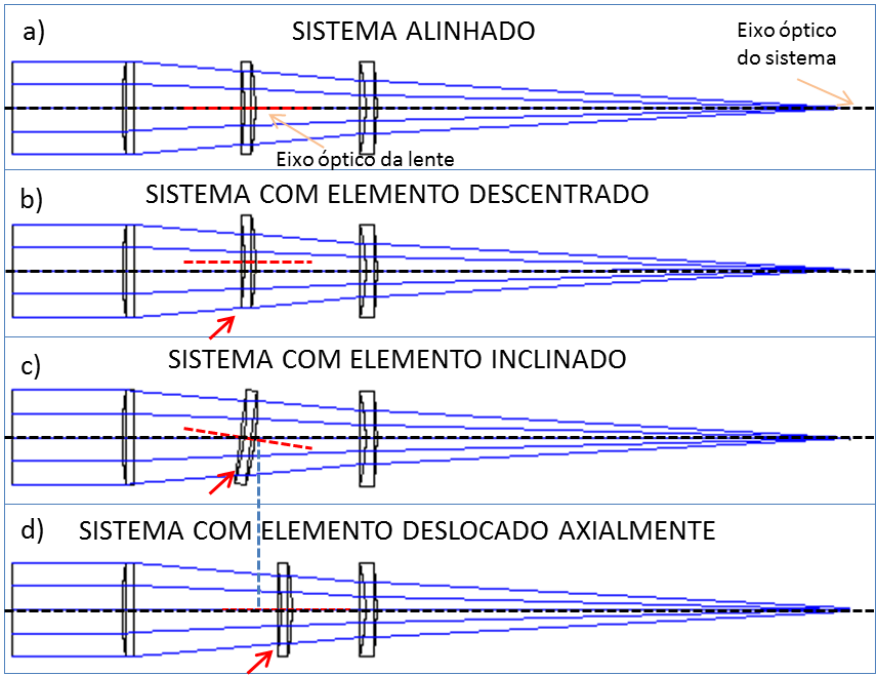

Figura 1 – Erros de alinhamento nos elementos ópticos; a) Sistema alinhado; b) Elemento descentralizado; c) Elemento inclinado; d) Elemento deslocado axialmente.

- DESLOCAMENTO AXIAL:
  Elementos deslocados ao longo do eixo óptico. Ocorrem desvios no valor da separação entre elementos (Figura 1, d).

  Em sistemas rotacionalmente simétricos, estes deslocamentos não causam perda da simetria axial e, portanto, não afetam as aberrações do sistema da mesma forma que descentralizações e inclinações afetam.

  Deslocamentos axiais alteram, principalmente, aberração esférica e a posição de foco.

As teorias de aberrações de onda aplicáveis a sistemas não axialmente simétricos, descritas neste capítulo, se aplicam apenas a descentralizações e inclinações nos elementos ópticos.

2.2 Aberrações em sistemas ópticos

Imagens formadas por sistemas ópticos reais apresentam defeitos causados por difração e aberrações geométricas. A melhor imagem possível, conhecida como imagem limitada por difração (*diffraction-limited*) ocorre quando as aberrações geométricas não estão presentes (17).

As aberrações podem ser divididas em aberrações cromáticas e monocromáticas:

- Aberrações cromáticas: ocorrem em sistemas refrativos devido à variação do índice de refração com o comprimento de onda.
- Aberrações monocromáticas: ocorrem em sistemas refrativos e reflexivos e não dependem do comprimento de onda. São as aberrações geométricas.

Neste trabalho, apenas as aberrações monocromáticas serão abordadas.

As aberrações presentes em sistemas ópticos podem ser calculadas por meio das aberrações de raios (*ray aberration*) ou das aberrações de onda (*wave aberration*) (36).

O caminho óptico de um raio é dado pela multiplicação do índice de refração do meio e o caminho geométrico percorrido pelo raio. Se um conjunto de raios parte de um objeto pontual, atravessa o sistema óptico e chega à pupila de saída, sendo que cada raio percorre um caminho óptico igual ao do raio principal[c], a frente de onda do sistema, para este ponto objeto, é dada pela superfície que conecta os pontos das extremidades destes raios.

Se a frente de onda é esférica, as aberrações de onda são nulas, os raios convergem para o seu centro de curvatura (no ponto imagem gaussiano) e uma imagem perfeita do objeto pontual é formada. As aberrações de frente de onda, para um objeto pontual, representam o desvio óptico (desvio do caminho geométrico multiplicado pelo índice de refração do meio) da sua frente de onda em relação a uma esfera na pupila de saída do sistema (esfera de referência gaussiana) (36).

[c] *Chief ray:* raio que passa pelo centro do *STOP* de abertura do sistema. Fonte: The Photonics Dictionary, Pittsfield: Laurin Publishing, p. D22, 2004.

A aberração de onda para um raio, em um ponto da esfera de referência que é atingido por este raio, é igual ao desvio óptico da frente de onda ao longo daquele raio em relação à frente de onda esférica gaussiana e representa a diferença de caminho óptico (*optical path difference* - OPD) deste raio particular em relação ao raio principal no percurso que parte do ponto objeto até a esfera de referência. Dessa forma, a aberração de onda do raio principal é nula (36).

Considerando que os caminhos ópticos dos raios que partem da esfera de referência até o ponto imagem gaussiano são iguais, a aberração de onda de um raio é também dada pela diferença entre seu caminho óptico e o do raio principal desde o ponto objeto até o ponto imagem gaussiano. A aberração de onda de um raio é positiva se ele percorrer um caminho óptico maior que o do raio principal para atingir a esfera de referência gaussiana (36).

As aberrações de raios para um objeto pontual, obtidas por meio do traçado de raios através do sistema óptico até o plano imagem, representam o deslocamento dos raios em relação ao centro de curvatura da esfera de referência (ponto imagem gaussiano), no plano imagem. Essas aberrações podem também ser derivadas das aberrações de onda e são proporcionais às derivadas das aberrações de onda com relação às coordenadas da pupila. As aberrações de raios não levam em consideração a difração da frente de onda na pupila de saída do sistema óptico (36).

Sendo $W(x,y)$ a aberração de onda em um ponto (*x, y, z*) na esfera de referência, n_i o índice de refração do espaço imagem, R' a distância entre a esfera de referência e o ponto $P''(x_i,y_i)$ (dado pela posição em que um raio cruza o plano imagem gaussiano), a aberração de raio $(x_i - x_g, y_i)$ no plano imagem gaussiano para um ponto objeto localizado no eixo x (sendo z a direção do eixo óptico) é dada por (36):

$$x_i - x_g = \frac{R'}{n_i}\frac{\partial W}{\partial x}\ , \quad y_i = \frac{R'}{n_i}\frac{\partial W}{\partial y}\ . \tag{1}$$

Considerando-se que $R' \cong R$, onde R é o raio de curvatura da esfera de referência (é a distância entre o plano da pupila de saída e o plano imagem gaussiano, muito maior que a extensão da distribuição dos raios no plano imagem), tem-se:

$$\left(x_i, y_i\right) = \frac{R}{n_i}\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right), \tag{2}$$

onde $\left(x_i, y_i\right)$ são as coordenadas no ponto P'' em relação ao ponto imagem gaussiano.

Considerando-se $W\left(x,y\right)$ a aberração de onda de um ponto $\left(x,y\right)$ projetado no plano da pupila de saída, em coordenadas polares, pode-se escrever:

$$\left(x,y\right) = r\left(\cos\theta, \operatorname{sen}\theta\right). \tag{3}$$

Assim:

$$x_i = \frac{R}{n_i}\left(\cos\theta \frac{\partial W}{\partial r} - \frac{\operatorname{sen}\theta}{r}\frac{\partial W}{\partial \theta}\right), \tag{4}$$

$$y_i = \frac{R}{n_i}\left(\operatorname{sen}\theta \frac{\partial W}{\partial r} + \frac{\cos\theta}{r}\frac{\partial W}{\partial \theta}\right). \tag{5}$$

Para um sistema com muitos elementos, as aberrações de onda são aditivas, ou seja, a aberração de onda de um raio para o sistema todo é igual à soma das aberrações de onda em cada elemento do sistema. O mesmo não acontece para as aberrações de raios, ou seja, as aberrações de um raio no plano imagem final não podem ser obtidas por meio da soma dos seus valores nos planos imagem de cada elemento.

Este capítulo apresenta uma breve revisão das aberrações existentes num sistema óptico com simetria rotacional.

Também será apresentada uma descrição da teoria de aberrações de onda (*Wave aberration theory*) para sistemas simétricos, para sistemas plano-simétricos e uma generalização para sistema não simétricos.

Essas abordagens são baseadas nos trabalhos de José Sasián (3) e Kevin Thompson (2).

2.2.1 Aberrações em sistemas ópticos com simetria rotacional

Em sistemas ópticos com simetria rotacional, a função de aberração de onda $W(\mathbf{H},\boldsymbol{\rho})$ é descrita por um somatório em série de potências de combinações dos vetores de campo ($\mathbf{H}$) e da pupila ($\boldsymbol{\rho}$) dado por (3):

$$W(\mathbf{H},\boldsymbol{\rho}) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} W_{2k+n,2m+n,n} (\mathbf{H}\cdot\mathbf{H})^{k} (\boldsymbol{\rho}\cdot\boldsymbol{\rho})^{m} (\mathbf{H}\cdot\boldsymbol{\rho})^{n}. \tag{6}$$

Ou, na sua forma escalar:

$$W(H,\rho,\varphi) = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} W_{2k+n,2m+n,n} H^{2k+n} \rho^{2m+n} \cos^{n}\varphi. \tag{7}$$

A constante $W_{2k+n,2m+n,n}$, é dada pela soma das contribuições de todas as superfícies ($_j$) do sistema:

$$W_{2k+n,2m+n,n} = \sum_{j} (W_{2k+n,2m+n,n})_{j}. \tag{8}$$

Esta constante consiste no fator peso de cada termo do somatório, e $2k+n$ determina a dependência da aberração com relação ao vetor de campo, $2m+n$ indica a dependência com relação ao vetor da pupila, e n indica a dependência em relação ao ângulo subentendido entre esses dois vetores.

Quando as potências de H e ρ são somadas, obtém-se a ordem da aberração, que é sempre par, pois é dada por $2(k+m+n)$. As nomenclaturas das aberrações são devidas às suas dependências com ρ e φ. A dependência com H pode ser considerada uma característica de cada termo de aberração e define a terminologia: constante, linear, quadrática etc dadas a elas.

Os termos de segunda ordem correspondem a alterações nas propriedades de primeira ordem do sistema e não são considerados aberrações. A Tabela 1 exibe esses termos.

Tabela 1 – Termos de segunda ordem da função de aberração de onda.

Forma escalar	Coeficiente	Nome
$W_{200}H^2$	W_{200}	*Piston* (quadrático)
$W_{020}\rho^2$	W_{020}	*Defocus*
$W_{111}H\rho\cos(\varphi)$	W_{111}	*Magnification* - ampliação

Fonte: Adaptada de SASIÁN (3, 37).

Enquanto *piston*[d] não apresenta efeitos na imagem, o termo denominado *defocus* afeta a posição da imagem axial, e o termo *magnification* afeta o tamanho da imagem (3).

Os termos de quarta ordem são conhecidos como aberrações primárias ou aberrações de Seidel, que descrevem as aberrações de baixa ordem que ocorrem em sistemas com simetria rotacional. Os cinco termos que constituem esse conjunto são apresentados na Tabela 2.

Tabela 2 – Termos de quarta ordem da função de aberração de onda.

Forma escalar	Coeficiente	Nome
$W_{040}\rho^4$	W_{040}	Aberração esférica
$W_{131}H\rho^3\cos(\varphi)$	W_{131}	Coma (Linear)
$W_{222}H^2\rho^2\cos^2(\varphi)$	W_{222}	Astigmatismo (Quadrático)
$W_{220}H^2\rho^2$	W_{220}	Curvatura de campo
$W_{311}H^3\rho\cos(\varphi)$	W_{311}	Distorção (Cúbica)

Fonte: Adaptada de SASIÁN (3, 37).

Os termos de sexta ordem são conhecidos como aberrações secundárias ou aberrações de Schwarzchild, e as de oitava ordem são as aberrações terciárias. Detalhes sobre estas aberrações podem ser encontrados em Sasián (37) e Hopkins (38).

Conforme descrito anteriormente neste capítulo, as aberrações de raios são obtidas por meio da derivada espacial das aberrações de onda; assim, a ordem das aberrações de raio (descritas no plano imagem) é sempre uma ordem menor que a das aberrações de onda (descritas na pupila de saída). Os termos de quarta ordem nas aberrações de onda correspondem aos de terceira ordem nas aberrações de raios (36, 39).

[d] *Piston* não afeta a qualidade da imagem de um objeto pontual pois representa um atraso ou avanço na propagação da frente de onda. Fonte: Sasián J. Introduction to aberrations in optical imaging systems. New York: Cambridge University Press, 2013. 261p.

2.2.2 Aberrações de Seidel

A seguir, serão descritas algumas das características de cada termo que compõe as aberrações de Seidel, ou aberrações de terceira ordem. As dependências com o campo descritas a seguir são válidas para sistemas com simetria rotacional. Mais detalhes podem ser encontrados em Sasián (3), Smith (17), Mahajan (36), Slyusarov (40).

- Aberração esférica

Esta aberração pode ser definida como sendo a variação da posição de foco com a abertura (17). Ela ocorre quando raios provenientes de um objeto pontual que passam por diferentes regiões (zonas) da pupila formam imagem em diferentes posições axiais, ou seja, raios mais próximos ao eixo óptico formam imagem em uma posição próxima à posição de foco paraxial, diferentemente de raios marginais (raios que atravessam a pupila em regiões próximas à borda) (Figura 2, A).

A imagem de um objeto pontual, formada por um sistema com aberração esférica é geralmente um ponto brilhante contornado por um halo de luz. Para objetos extensos, o efeito desta aberração na imagem é a redução do contraste e da clareza dos detalhes (17).

A aberração esférica possui dependência de ρ^4 com a posição na pupila e não possui dependência com o campo (H); assim, seu valor é o mesmo em todo o campo.

Mesmo para um objeto no eixo óptico, quando as outras aberrações monocromáticas desaparecem, a aberração esférica ainda pode existir (40).

- Coma

O coma pode ser definido como a variação da magnificação com a abertura e ocorre quando raios vindos de diferentes zonas na pupila formam imagem em diferentes posições (alturas) do campo no plano imagem (17) (Figura 2, B).

O efeito causado pelo coma é que a imagem de um objeto pontual possui a forma de um cometa.

A dependência do coma com a posição na pupila é dada por $\rho^3 \cos\varphi$ e é linear com o campo, sendo nula sobre o eixo óptico.

Devido a essa dependência linear, seu efeito em ângulos pequenos do campo de visada é maior que o de aberrações que possuem dependência quadrática ou cúbica com o campo. No caso de telescópios com dois espelhos do tipo Cassegrain, esta é a aberração limitante do desempenho mais importante (5). Já telescópios denominados Ritchey-Chrétien são corrigidos tanto para aberração esférica como coma em todo o campo; neste caso, a aberração limitante é o astigmatismo (20).

A chamada Violação à condição de seno de Abbe (*Offense against the Abbe sine condition - OSC*) fornece uma indicação da quantidade de coma existente no sistema em regiões próximas ao eixo óptico (17).

- Astigmatismo

O astigmatismo ocorre quando os raios tangenciais e sagitais formam imagem em diferentes posições axiais (Figura 2, D).

Pode-se definir raios tangenciais como raios contidos no plano tangencial. Para um objeto posicionado no eixo *y*, o plano tangencial é o *y-z*, e o sagital é o *x-z*.

A imagem de um objeto pontual, formada pelos raios no plano tangencial, será uma linha localizada no plano sagital. Da mesma forma, a imagem de um objeto pontual formada pelos raios no plano sagital será uma linha localizada no plano imagem tangencial. O astigmatismo ocorre quando as imagens sagital e tangencial não coincidem.

A imagem de um objeto pontual formada por um sistema óptico astigmático será constituída por duas linhas distintas.

Entre os focos astigmáticos, a imagem será um “borrão” de luz elíptico ou circular (17).

O astigmatismo possui dependência $\rho^2\varphi^2$ com a posição na pupila e dependência quadrática com a posição no campo; assim, se uma lente é bem feita, não haverá astigmatismo na imagem de um objeto pontual sobre o eixo óptico.

- Curvatura de campo

A curvatura de campo ocorre porque o plano imagem, ou melhor, a superfície de imagem mais nítida na verdade não é plana. Para um determinado ponto do campo, a curvatura de campo ocorre como um erro de foco. Assim, esta aberração pode ser entendida como um *defocus* com dependência específica do campo (Figura 2, D).

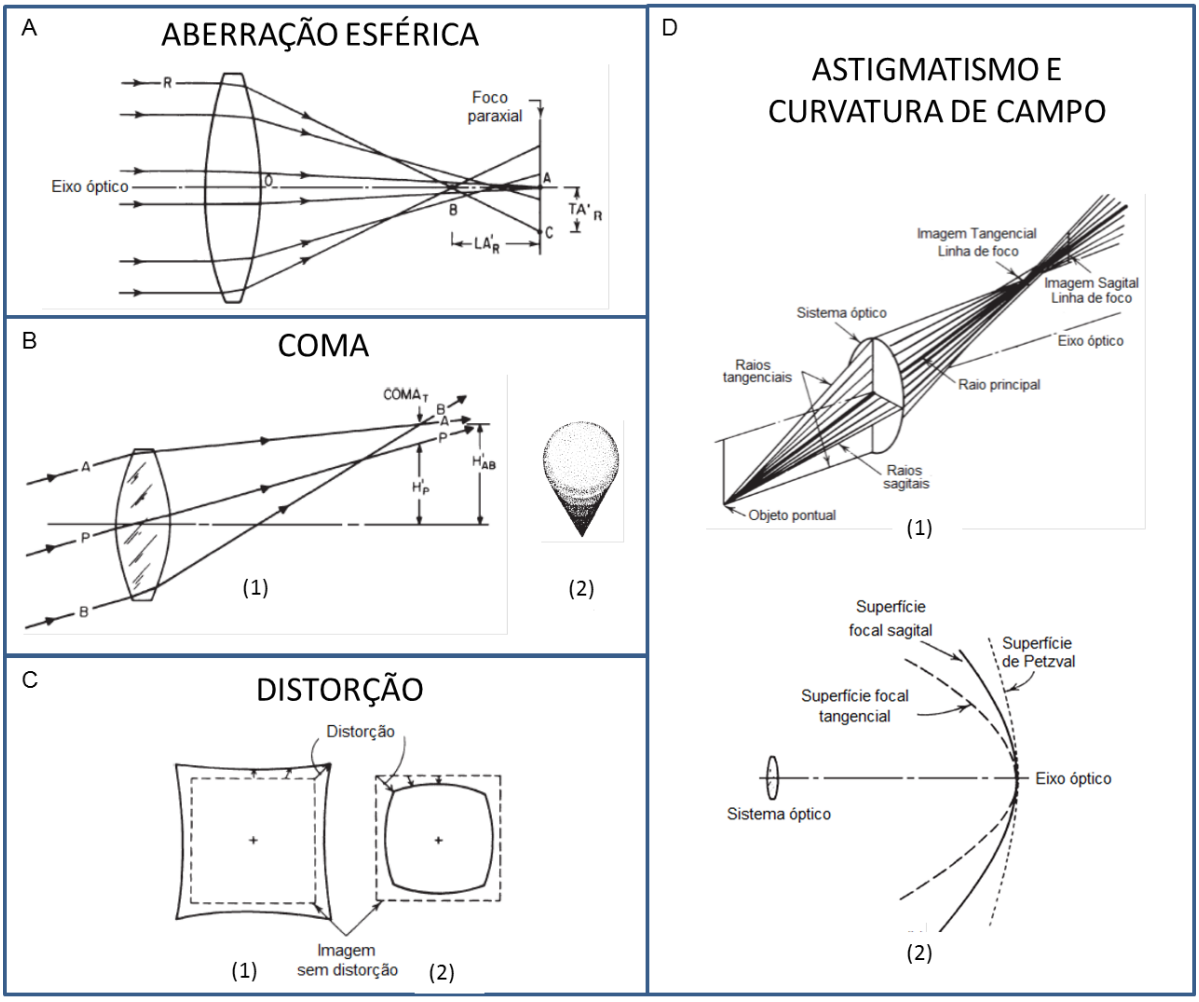

Figura 2 – Aberrações de Seidel. A) Lente simples com aberração esférica onde raios que atravessam a lente em regiões mais afastadas do eixo óptico formam imagem em pontos mais próximos da lente; B) Lente simples com coma: 1) Raios que passam por regiões mais afastadas do eixo óptico formam imagem em alturas diferentes no plano imagem dos raios que passam pelo eixo; 2) Imagem de um objeto pontual formada por um sistema que apresenta coma. C) Comparação de imagens de sistemas com Distorção: 1) Distorção *pincushion* ou positiva; 2) Distorção *barrel* (barril) ou negativa. D) Astigmatismo e curvatura de campo: 1) Sistema com astigmatismo; 2) Ilustração das superfícies imagem sagital e tangencial posicionadas em relação à superfície de Petzval para um sistema com astigmatismo Fonte: Adaptada de SMITH (17).

Todo sistema óptico tem uma curvatura de campo “básica” chamada de curvatura de Petzval. Quando o sistema é livre de astigmatismo, as superfícies imagem sagital e tangencial coincidem com a superfície de Petzval. Quando há astigmatismo primário

no sistema, a superfície imagem tangencial está localizada três vezes mais distante da superfície de Petzval que a superfície imagem sagital, ambas deslocadas para a mesma direção (17).

Curvatura de campo possui dependência ρ^2 com a pupila e dependência quadrática com o campo.

- Distorção

Distorção não degrada a imagem, porém altera a posição onde a imagem de um objeto deveria ser formada (Figura 2, C). Pode também ser considerada como um erro de magnificação com dependência com o campo.

Possui dependência $\rho\cos(\varphi)$ com a pupila e dependência cúbica com o campo.

2.2.3 Sistemas ópticos não rotacionalmente simétricos ou desalinhados

Em sistemas ópticos axialmente simétricos, as aberrações de terceira ordem que podem estar presentes são as aberrações de Seidel apresentadas anteriormente. Quando a simetria axial é quebrada de forma planejada ou devido a erros de alinhamento, surgem termos de aberrações com diferentes dependências com o campo.

Em sua tese, Buchroeder (19), em 1976, estudou as aberrações em sistemas não rotacionalmente simétricos, e seus estudos resultaram na hipótese de que o campo de aberração no plano imagem destes sistemas não rotacionalmente simétricos ou desalinhados é ainda dado pela soma das contribuições individuais de cada superfície, embora as contribuições individuais não apresentem mais um centro comum sobre o eixo óptico.

Thompson (2) propôs que, em sistemas sem simetria rotacional, nenhuma aberração nova é criada, mas surgem novas dependências com o campo para as aberrações já conhecidas. Estas novas dependências possuem ordens mais baixas que as presentes nos sistemas simétricos axialmente (6, 20).

Shack e Thompson explicaram que, quando há elementos desalinhados no sistema, as contribuições individuais das suas superfícies permanecem rotacionalmente simétricas em torno de algum ponto no campo. Para superfícies asféricas, devido ao seu eixo único, o deslocamento do centro do campo é diferente para a porção esférica e a porção asférica que compõem a superfície, mas, para ambas as porções, a natureza simétrica da aberração em torno de seu centro deslocado será mantida. O que ocorre, porém, é que a combinação desses efeitos gera dependências complexas com o campo (5, 20, 23, 41).

A seguir, serão apresentadas duas teorias propostas para o entendimento e cálculo das aberrações em sistemas não rotacionalmente simétricos. A primeira delas, aplicável a sistemas plano-simétricos, representa uma teoria unidimensional. A segunda corresponde a uma teoria bidimensional aplicável aos desalinhamentos: inclinação e descentralização.

2.2.3.1 Sistemas ópticos plano-simétricos

Um sistema plano-simétrico ou bilateral apresenta um plano de simetria, sendo metade do sistema a imagem espelhada do outro. Sistemas com simetria axial ou sistemas duplo-plano-simétrico são também plano-simétricos.

Sistemas plano-simétricos representam um tipo particular de sistema não rotacionalmente simétrico, ou seja, sistemas que não apresentam um eixo de simetria rotacional. Nesses sistemas, os componentes ópticos podem estar inclinados no plano de simetria.

O conhecimento desse tipo de sistema é de grande relevância no entendimento dos efeitos dos erros de fabricação e montagem em sistema ópticos inicialmente projetados para serem axialmente simétricos. Pode também ser aplicado no projeto de sistemas reflexivos sem obscuração central (3).

Sistemas que apresentam um plano de simetria bilateral são interessantes, pois são, de certa forma, simples de serem descritos, mas ainda apresentam as principais diferenças entre sistemas sem e com simetria axial.

Esta revisão se baseia no trabalho de Sasián (3), que representa uma extensão da teoria de aberração de onda para sistemas rotacionalmente simétricos compostos por superfícies esféricas ou ligeiramente asféricas. Em sua formulação, a função de aberração de sistemas plano-simétricos é descrita e fornece a posição e o tamanho da imagem e os defeitos presentes na imagem. Os campos de aberrações também são discutidos com detalhes.

As definições adotadas pelo autor e a descrição detalhada dos coeficientes das aberrações de onda são encontradas no apêndice A.

A função de aberração representa as deformações da frente de onda ao longo do campo de visada e da abertura do sistema. Para um sistema plano-simétrico é dada por:

$$W(\mathbf{H},\boldsymbol{\rho}) = \sum_{k,m,n,p,q}^{\infty} W_{2k+n+p,2m+n+q,n,p,q} (\mathbf{H}\cdot\mathbf{H})^k (\boldsymbol{\rho}\cdot\boldsymbol{\rho})^m (\mathbf{H}\cdot\boldsymbol{\rho})^n (\mathbf{i}\cdot\mathbf{H})^p (\mathbf{i}\cdot\boldsymbol{\rho})^q. \quad (9)$$

Na equação (9), $W_{2k+n+p,2m+n+q,n,p,q}$ é o coeficiente de uma determinada aberração definida pelos inteiros *k, m, n, p* e *q*.

As potências de H, ρ, φ, α e β são dadas respectivamente por $2k+n+p$, $2m+n+p$, n, p e q.

A equação (9) apresentada anteriormente é uma generalização da função de aberração de um sistema axialmente simétrico. Quando *p* e *q* são nulos, a função de aberração de um sistema axialmente simétrico, equação (6), é obtida.

A seguir, é apresentada uma tabela onde as aberrações do sistema são divididas em grupos. Os três primeiros grupos são apresentados e representam o grau de aproximação com a deformação da frente de onda. Tais grupos são obtidos quando a soma dos números inteiros é igual a 0, 1, 2, 3, etc. A terminologia apresentada na Tabela 3 (constante, linear etc) é dada com relação à potência do vetor de campo (3).

Além dos grupos, as aberrações são organizadas em subgrupos que contêm as aberrações características de sistemas axialmente simétricos, sistemas com dois planos de simetria e sistemas com um plano de simetria.

O primeiro grupo é composto pelo termo *piston*. Este termo representa uma mudança de fase uniforme ao longo do campo de visada que não afeta a qualidade nem a posição da imagem, já que não há dependência com a abertura.

O segundo grupo, já descrito anteriormente, é composto por *defocus* (desfocalização), que representa um desvio longitudinal da posição da imagem ideal; *magnification* ou magnificação (alteração de escala), que representa um desvio do tamanho da imagem ideal; e *piston* quadrático (deslocamento de campo ou *boresight error*), que causa um deslocamento (o*ffset*) do centro do campo de visada. Entretanto, como por definição o OAR[e] cruza o centro da imagem, este termo é nulo. Pode-se definir a posição e tamanho da imagem de modo que esses três parâmetros sejam nulos; isso implica que a posição e tamanho da imagem são conhecidos. Dessa forma, o segundo grupo define a posição e tamanho da imagem.

O terceiro grupo é composto pelas aberrações que devem ser balanceadas para a obtenção de uma imagem com boa qualidade. Além das aberrações de Seidel (terceiro subgrupo do terceiro grupo, que representam as aberrações de um sistema axialmente simétrico e que já foram discutidas anteriormente), existem mais sete outras aberrações que devem ser consideradas no desenvolvimento de um sistema plano-simétrico. O segundo subgrupo exibe as aberrações de um sistema plano-simétrico. As aberrações de um sistema com dois planos de simetria são apresentadas no primeiro subgrupo do terceiro grupo.

As deformações da frente de onda representadas pelas aberrações de um sistema plano-simétrico não apresentam novas formas além daquelas que já são conhecidas para sistemas com simetria rotacional. O que é diferente, em relação a um sistema axialmente simétrico, é a distribuição da aberração ao longo do campo de visada. Um bom exemplo é o ponto de correção da aberração, que é o ponto onde a aberração tem valor nulo e ela é simétrica em relação a este ponto; pode ter diferentes localizações para as diferentes aberrações. Outra diferença é que a imagem pode ser anamórfica, e o plano imagem pode ser inclinado em relação ao OAR.

[e] OAR (*Optical axis ray*), definido no apêndice A.

Tabela 3 – Grupos de aberrações em sistemas plano-simétricos, de acordo com a definição de Sasián (3).

Forma vetorial	Forma escalar	Nome
	Primeiro Grupo	
W_{00000}	W_{00000}	*Piston* constante
	Segundo Grupo	
$W_{01001}(\mathbf{i}\cdot\boldsymbol{\rho})$	$W_{01001}\,\rho\cos(\beta)$	Deslocamento de campo
$W_{10010}(\mathbf{i}\cdot\mathbf{H})$	$W_{10010}\,H\cos(\alpha)$	*Piston* linear
$W_{02000}(\boldsymbol{\rho}\cdot\boldsymbol{\rho})$	$W_{02000}\rho^2$	*Defocus*
$W_{11100}(\mathbf{H}\cdot\boldsymbol{\rho})$	$W_{11100}H\rho\cos(\varphi)$	Magnificação
$W_{20000}(\mathbf{H}\cdot\mathbf{H})$	$W_{20000}H^2$	*Piston* quadrático
	Terceiro Grupo	
$W_{02002}(\mathbf{i}\cdot\boldsymbol{\rho})^2$	$W_{02002}\rho^2\cos^2(\beta)$	Astigmatismo constante
$W_{11011}(\mathbf{i}\cdot\mathbf{H})(\mathbf{i}\cdot\boldsymbol{\rho})$	$W_{11011}H\rho\cos(\alpha)\cos(\beta)$	Anamorfismo
$W_{20020}(\mathbf{i}\cdot\mathbf{H})^2$	$W_{20020}H^2\cos^2(\alpha)$	*Piston* quadrático
$W_{03001}(\mathbf{i}\cdot\boldsymbol{\rho})(\boldsymbol{\rho}\cdot\boldsymbol{\rho})$	$W_{03001}\rho^3\cos(\beta)$	Coma constante
$W_{12101}(\mathbf{i}\cdot\boldsymbol{\rho})(\mathbf{H}\cdot\boldsymbol{\rho})$	$W_{12101}H\rho^2\cos(\varphi)\cos(\beta)$	Astigmatismo linear
$W_{12010}(\mathbf{i}\cdot\mathbf{H})(\boldsymbol{\rho}\cdot\boldsymbol{\rho})$	$W_{12010}H\rho^2\cos(\alpha)$	Inclinação de campo
$W_{21001}(\mathbf{i}\cdot\boldsymbol{\rho})(\mathbf{H}\cdot\mathbf{H})$	$W_{21001}H^2\rho\cos(\beta)$	Distorção quadrática I
$W_{21110}(\mathbf{i}\cdot\mathbf{H})(\mathbf{H}\cdot\boldsymbol{\rho})$	$W_{21110}H^2\rho\cos(\varphi)\cos(\alpha)$	Distorção quadrática II
$W_{30010}(\mathbf{i}\cdot\mathbf{H})(\mathbf{H}\cdot\mathbf{H})$	$W_{30010}H^3\cos(\alpha)$	*Piston* cúbico
$W_{04000}(\boldsymbol{\rho}\cdot\boldsymbol{\rho})^2$	$W_{04000}\rho^4$	Aberração esférica
$W_{13100}(\mathbf{H}\cdot\boldsymbol{\rho})(\boldsymbol{\rho}\cdot\boldsymbol{\rho})$	$W_{13100}H\rho^3\cos(\varphi)$	Coma linear
$W_{22200}(\mathbf{H}\cdot\boldsymbol{\rho})^2$	$W_{22200}H^2\rho^2\cos^2(\varphi)$	Astigmatismo quadrático
$W_{22000}(\mathbf{H}\cdot\mathbf{H})(\boldsymbol{\rho}\cdot\boldsymbol{\rho})$	$W_{22000}H^2\rho^2$	Curvatura de campo
$W_{31100}(\mathbf{H}\cdot\mathbf{H})(\mathbf{H}\cdot\boldsymbol{\rho})$	$W_{31100}H^3\rho\cos(\varphi)$	Distorção cúbica
$W_{40000}(\mathbf{H}\cdot\mathbf{H})^2$	$W_{40000}H^4$	*Piston* quarta ordem

Fonte: Adaptada de SASIÁN (3).

A distribuição da aberração ao longo do campo de visada muda porque novos termos de aberrações são gerados quando o sistema deixa de ter simetria rotacional. Tais termos sempre possuem ordens (relativas ao vetor do campo) menores que os termos que podem estar presentes em sistemas rotacionalmente simétricos. Assim, coma, que possui dependência linear com o campo em sistemas

rotacionalmente simétricos terá um termo adicional que possui dependência constante com o campo em sistemas não rotacionalmente simétricos. Para astigmatismo, inicialmente com dependência quadrática com o campo, podem surgir astigmatismos linear e constante.

A aberração chamada de astigmatismo linear (pois é linear com o campo) ou *linear astigmatism* é uma medida da diferença entre as inclinações do plano imagem sagital e tangencial.

No terceiro grupo, existem dois termos de distorção quadrática: o termo I produz um deslocamento da imagem no plano de simetria quando um círculo é imageado, que é proporcional ao quadrado do raio do círculo; o termo II é o termo principal da chamada *Keystone distortion*.

Dessa forma, a separação em subgrupos de acordo com a simetria do sistema leva a uma estratégia de correção de sistemas plano-simétricos que pode ser dada pela correção subsequente dos diferentes subgrupos, iniciando-se pela correção das aberrações de Seidel por meio de técnicas padrões e, em seguida, corrigindo ou balanceando as aberrações de sistemas com um ou dois planos de simetria.

A correção de um dado sistema pode ser feita por meio da correção de cada subgrupo utilizando-se as variáveis adequadas para cada simetria, ou seja, as aberrações de Seidel podem ser influenciadas pelas asfericidades de superfícies que são axialmente simétricas; as aberrações de sistemas com dois planos de simetria podem ser influenciadas pelas asfericidades que possuem a mesma simetria, e as de sistemas plano-simétricos podem ser influenciadas por inclinação de superfícies e asfericidades plano-simétricas.

Campos de aberrações:

Campos de aberrações são associados a termos de aberrações que apresentam uma determinada dependência com a abertura. No terceiro grupo, existem seis campos correspondentes às aberrações de Seidel (incluindo um termo *piston*). No caso de sistemas que não possuem simetria axial, os campos de aberrações tornam possível haver mais um ponto de correção (*point node*) ou uma linha de correção (*line node*) em que uma aberração particular é nula.

Neste grupo, para o caso da aberração esférica, constante ao longo de todo o campo, não aparece nenhum termo novo.

No caso do coma de terceira ordem, para sistemas axialmente simétricos, essa aberração é corrigida para todo o campo ou apresenta uma dependência linear com ele. Em sistemas com simetria linear, o coma pode ser constante e diferente de zero ao longo de todo o campo. Quando há coma constante e linearmente dependente do campo, continua existindo apenas um nodo, porém não mais no centro do campo.

Para o astigmatismo, em sistemas plano-simétricos, existem: astigmatismo constante, linear e quadrático com o campo. Neste caso, dois nodos podem coexistir levando ao astigmatismo binodal (2-3, 5). Estes dois nodos coincidem e são localizados no centro do campo para sistemas axialmente simétricos.

A curvatura de campo é formada por inclinação e curvatura do campo, possui um nodo no centro do campo e pode conter um nodo linear ou circular.

O campo da distorção é formado por distorção anamórfica, quadrática I e II e distorção cúbica. Possui um nodo central e pode haver outros nodos ou linha de nodos.

Sistemas com simetria bilateral podem apresentar nodos ao longo do plano de simetria ou em linhas perpendiculares a estes planos, e sua distribuição deve ter a mesma simetria bilateral do sistema (3).

Coeficiente de aberração:

Os coeficientes de aberração representam o máximo desvio da frente de onda em relação à frente de onda esférica básica. Quando esses coeficientes são superpostos e somados, a soma resulta na deformação total medida na pupila de saída e por todo o campo de visada.

Conforme pode ser observado no apêndice A, os novos termos, que não aparecem em sistemas axialmente simétricos, possuem dependência com o ângulo de inclinação da superfície (I) em relação ao OAR. Desta forma, estes termos valem para sistemas que apresentam elementos inclinados em relação ao OAR. Os termos adicionais, quando o elemento é ligeiramente asférico, não apresentam dependências com a constante cônica da superfície.

No caso de elementos descentralizados em relação ao OAR, as equações para a contribuição esférica podem ser adaptadas. Neste caso, a contribuição asférica deve

ser calculada e terá dependência com a constante cônica. Maiores informações podem ser encontradas no capítulo 3.

2.2.3.2 Sistemas ópticos sem simetria rotacional

Nesta seção, será apresentada uma segunda abordagem para o estudo e cálculo dos efeitos de descentralização e inclinação em sistemas ópticos. Esta abordagem foi proposta por Kevin P. Thompson em sua tese de doutorado (2) e é conhecida como Teoria de Aberração de Onda Nodal (*Nodal Aberration Theory*).

O trabalho de Thompson baseou-se em duas hipóteses (ou conceitos) fundamentais:

Uma das hipóteses originou-se no trabalho de Buchroeder (19) e está relacionada com o comportamento dos campos de aberrações convencionais em sistemas ópticos perturbados. Com base em duas propriedades bem estabelecidas dos campos de aberrações em sistemas axialmente simétricos:

1) O campo de aberração no plano imagem é a soma das contribuições das aberrações de cada superfície do sistema;
2) A contribuição de uma superfície ao campo de aberração está centrada ao longo de uma linha que conecta os centros das pupilas para aquela superfície com o centro de curvatura dela.

Buchroeder estabeleceu a premissa: em sistemas ópticos perturbados, o campo de aberração no plano imagem é ainda dado pela soma das contribuições individuais de cada superfície do sistema, porém, estas contribuições não possuem mais um centro comum no plano imagem.

A premissa de Buchroeder estabelece o conceito de descentralização dos campos de aberrações convencionais no plano imagem, que possibilita ser feita a descrição do comportamento de um sistema perturbado utilizando-se a mesma terminologia usada para sistemas alinhados (2).

O segundo conceito fundamental adotado por Thompson para o desenvolvimento de seu trabalho originou-se do trabalho desenvolvido por Roland Shack. Shack

estabeleceu uma expressão vetorial para a expansão da aberração de onda usando uma multiplicação vetorial desenvolvida por ele mesmo. Esta expansão possibilita o tratamento de sistemas com perturbações aleatórias, não meridionais, não coplanares, sem causar problemas de notação.

A operação vetorial desenvolvida por Shack, denominada multiplicação vetorial, é diferente das operações produto escalar e produto vetorial, e nenhum sinal entre os vetores é usado para sua representação.

Os coeficientes das aberrações em uma superfície do sistema não são afetados por perturbações. Isso ocorre porque as propriedades paraxiais não são afetadas por inclinação e descentralização. O que ocorre na presença desses desalinhamentos é que o comportamento das aberrações no campo, ou seja, a dependência com o campo das aberrações, é modificado quando elementos descentralizados e inclinados estão presentes.

Um termo particular de aberração possui o seu centro de simetria deslocado do centro do plano imagem e, em muitos casos, um dado termo de aberração desenvolverá múltiplos zeros com simetria específica no plano imagem.

O estudo do comportamento dos zeros dos termos de aberrações possibilita um amplo entendimento do efeito de perturbações em sistemas ópticos e é a base dos estudos desenvolvidos por Thompson.

Definições:

- *reference axis* ou eixo de referência (RA): é o eixo definido no sistema perturbado dado pelo eixo mecânico do sistema em seu estado alinhado e não sofre alterações devido a perturbações. É o eixo de referência para a medida de inclinações e descentralizações, descentralizações do objeto, da imagem e das pupilas e inclinações do plano imagem gaussiano;
- *local axis* ou eixo local (LA): é definido, para uma dada superfície, como a linha que contém o vértice e o centro de curvatura da superfície. O ângulo entre o LA e o RA determina a inclinação da superfície. A descentralização é dada pela localização do vértice da superfície em relação ao RA;

- *optical axis ray* ou raio do eixo óptico (OAR): é o raio que passa pelo centro dos planos objeto e imagem, e pelos centros das pupilas para todas as superfícies.

Toda superfície ($_j$) de um sistema tem associada a ela um plano objeto (O_j) e uma pupila de entrada (E_j), sendo eles reais ou imaginários. A localização de O_j e E_j depende dos poderes das superfícies precedentes a $_j$. Estão também associados a esta superfície o plano imagem (O'_j) e a pupila de saída (E'_j), e estes se tornam o plano objeto ($O_{j+1} = O'_j$) e a pupila de entrada ($E_{j+1} = E'_j$) da superfície seguinte ($j+1$).

A contribuição de uma dada superfície para o campo de aberração está centrada na linha que conecta o centro da pupila daquela superfície com o seu centro de curvatura. A localização do centro de curvatura é, por sua vez, determinada pelas inclinações e descentralizações na superfície.

Em um sistema óptico perturbado, ou seja, com erros de alinhamento, o vetor de campo efetivo (*effective field height*) $\mathbf{H}_{Aj}$ substitui o vetor de campo $\mathbf{H}$ na equação de aberração de onda, equação (6), resultando em:

$$W\left(\mathbf{H},\boldsymbol{\rho},\boldsymbol{\sigma}_j\right) = \sum_j \sum_{k,m,n}^{\infty} \left(W_{2k+n,2m+n,n}\right)_j \left(\mathbf{H}_{Aj}\cdot\mathbf{H}_{Aj}\right)^k (\boldsymbol{\rho}\cdot\boldsymbol{\rho})^m (\mathbf{H}_{Aj}\cdot\boldsymbol{\rho})^n \quad , \tag{10}$$

onde:

$$\mathbf{H}_{Aj} = \mathbf{H} - \boldsymbol{\sigma}_j \ , \tag{11}$$

onde o vetor $\boldsymbol{\sigma}_j$ no plano imagem representa a localização do centro de simetria rotacional para uma superfície perturbada $_j$ em relação ao eixo óptico. Ele indica o ponto a partir do qual é medido o vetor de altura do campo efetivo. A localização do centro do campo de aberração pode ser calculada em um software de projeto óptico usando-se traçado de raios reais (*real-ray-trace data*) (23). A Figura 3 ilustra os vetores citados.

Para que a teoria de aberrações no plano imagem seja definida, é necessário definir um centro comum no plano imagem. Este ponto é definido como sendo o centro do plano imagem gaussiano, localizado sobre o OAR.

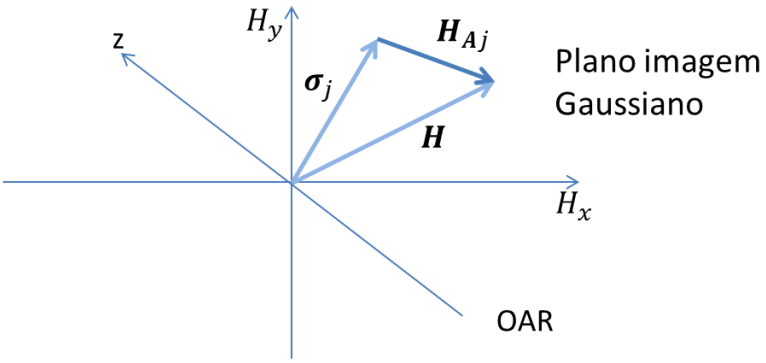

Figura 3 – Representação do vetor de campo efetivo. Fonte: THOMPSON (2).

Conforme já mencionado anteriormente, os coeficientes $\left(W_{2k+n,2m+n,n}\right)_j$ não mudam na presença de erros de alinhamento no sistema, pois dependem somente de quantidades paraxiais.

A equação apresentada anteriormente, com uma única modificação em relação à equação de aberração de onda obtida pela substituição de $\mathbf{H}$ por $\mathbf{H}_{Aj}$, é suficiente para abranger todas as modificações causadas por descentralização e inclinação. A equação para as aberrações de terceira ordem é dada por:

$$\begin{aligned} W = {} & \Delta W_{20}(\boldsymbol{\rho}\cdot\boldsymbol{\rho}) + \Delta W_{11}(\mathbf{H}\cdot\boldsymbol{\rho}) + \sum_j W_{040j}(\boldsymbol{\rho}\cdot\boldsymbol{\rho})^2 + \\ & \sum_j W_{131j}\left[(\mathbf{H}-\boldsymbol{\sigma}_j)\cdot\boldsymbol{\rho}\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho}) + \sum_j W_{222j}\left[(\mathbf{H}-\boldsymbol{\sigma}_j)\cdot\boldsymbol{\rho}\right]^2 + \\ & \sum_j W_{220j}\left[(\mathbf{H}-\boldsymbol{\sigma}_j)\cdot(\mathbf{H}-\boldsymbol{\sigma}_j)\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho}) + \\ & \sum_j W_{311j}[(\mathbf{H}-\boldsymbol{\sigma}_j)\cdot(\mathbf{H}-\boldsymbol{\sigma}_j)][(\mathbf{H}-\boldsymbol{\sigma}_j)\cdot\boldsymbol{\rho}]. \end{aligned} \tag{12}$$

Em sua tese, Thompson desenvolveu os termos de aberração de terceira ordem com detalhes. A seguir é apresentado um resumo dos resultados obtidos por ele. Nenhuma aberração nova é introduzida.

Aberração esférica: Dada pelo termo:

$$W = \sum_j W_{040j}(\boldsymbol{\rho}\cdot\boldsymbol{\rho})^2 \ . \tag{13}$$

Esta aberração não possui dependência com o campo $\mathbf{H}$ nem com $\boldsymbol{\sigma}_j$ e é, portanto, inerte aos efeitos de erros de alinhamento (descentralização e inclinação). Podem

ocorrer efeitos de ordens mais altas devido à inclinação dos planos objeto e imagem (2).

Coma: Dada pelo termo:

$$W = \sum_j W_{131j}\left[\left(\mathbf{H}-\boldsymbol{\sigma}_j\right)\cdot\boldsymbol{\rho}\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho}), \tag{14}$$

que pode ser reescrito como:

$$W = \left[\left(\left(\sum_j W_{131j}\mathbf{H}\right)-\left(\sum_j W_{131j}\boldsymbol{\sigma}_j\right)\right)\cdot\boldsymbol{\rho}\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho}), \tag{15}$$

onde o coeficiente do primeiro termo corresponde ao coeficiente do coma para sistemas alinhados. O segundo termo corresponde à soma das contribuições dos vetores de deslocamento de cada superfície no plano imagem, ponderados pela contribuição de cada superfície. O vetor resultante normalizado é:

$$\mathbf{a}_{131} = \frac{\mathbf{A}_{131}}{W_{131}} = \frac{\sum_j W_{131j}\boldsymbol{\sigma}_j}{W_{131}}, \tag{16}$$

resultando em uma equação para o campo do coma de terceira ordem em um sistema perturbado escrito em termos de quantidades no plano focal:

$$W = W_{131j}\left[\left(\mathbf{H}-\mathbf{a}_{131}\right)\cdot\boldsymbol{\rho}\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho})\,. \tag{17}$$

O efeito causado por perturbações é o deslocamento do campo do coma do sistema alinhado para um ponto localizado pelo vetor $\mathbf{a}_{131}$ (Figura 4,a).

Astigmatismo e curvatura de campo: Dados por:

$$W = \Delta W_{20}\left(\boldsymbol{\rho}\cdot\boldsymbol{\rho}\right) + \sum_j W_{220j}\left[\left(\mathbf{H}-\boldsymbol{\sigma}_j\right)\cdot\left(\mathbf{H}-\boldsymbol{\sigma}_j\right)\right](\boldsymbol{\rho}\cdot\boldsymbol{\rho}) + \sum_j W_{222j}\left[\left(\mathbf{H}-\boldsymbol{\sigma}_j\right)\cdot\boldsymbol{\rho}\right]^2. \tag{18}$$

Thompson define o astigmatismo em relação à superfície focal astigmática mediana (*medial astigmatic focal surface*) que corresponde, no caso do astigmatismo, à superfície de melhor foco, ou seja, à superfície que resulta na menor frente de onda RMS. Esta superfície é dada pelo coeficiente $W_{220M} = W_{220} + \frac{1}{2}W_{222}$.

Conforme pode ser observado com detalhes na tese de Thompson (2), a equação anterior resulta na seguinte equação para o astigmatismo, definido em função da superfície mediana em um sistema perturbado:

$$W = \tfrac{1}{2} W_{222} \left[\mathbf{H}_{222}^2 + \mathbf{b}_{222}^2 \right] \cdot \boldsymbol{\rho}^2 \ , \tag{19}$$

onde:

$$\mathbf{H}_{222}^2 = (\mathbf{H} - \mathbf{a}_{222})^2 \ . \tag{20}$$

Os pontos no campo onde o astigmatismo é zero podem ser encontrados por:

$$0 = (\mathbf{H} - \mathbf{a}_{222})^2 + \mathbf{b}_{222}^2 \ . \tag{21}$$

Seguindo as regras da multiplicação vetorial proposta por Shack, tem-se:

$$\mathbf{H} = \mathbf{a}_{222} \pm \{ -\mathbf{b}_{222}^2 \}^{1/2} \ . \tag{22}$$

Que resulta em duas raízes, ou seja, dois pontos no campo onde esta aberração é zero:

$$\mathbf{H} = \mathbf{a}_{222} \pm i\mathbf{b}_{222} \ . \tag{23}$$

A equação (23) indica que em um sistema perturbado, o astigmatismo em relação a uma superfície mediana (*medial surface*) é zero em dois pontos no campo (Figura 4, b) que corresponde ao astigmatismo binodal (*binodal astigmatism*). No caso do sistema alinhado, quando o sistema não é corrigido para astigmatismo, o fato de o astigmatismo apresentar apenas um ponto de zero sobre o eixo é apenas um caso particular em que os dois nodos coincidem no ponto central do campo.

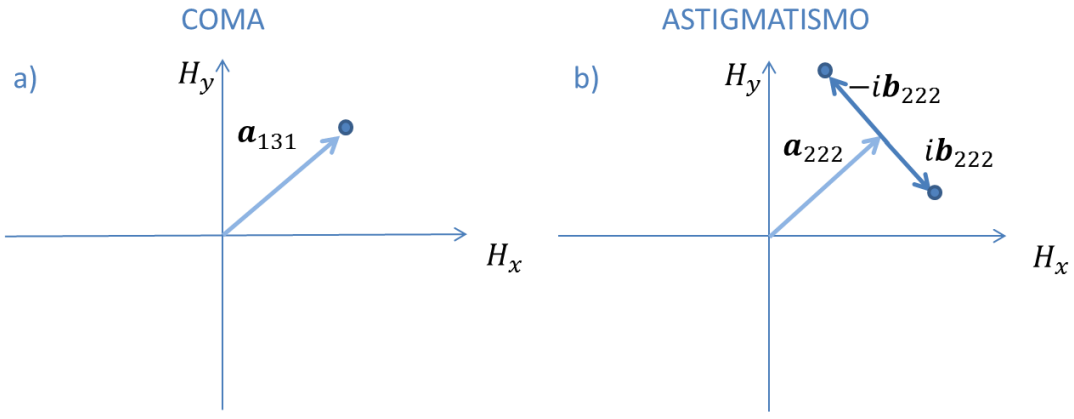

Figura 4 – Pontos onde a aberração é zero: a) Coma; b) Astigmatismo. Fonte: THOMPSON (2).

Distorção: Dada por:

$$W = \sum_j W_{311j} [(\mathbf{H} - \boldsymbol{\sigma}_j) \cdot (\mathbf{H} - \boldsymbol{\sigma}_j)][(\mathbf{H} - \boldsymbol{\sigma}_j) \cdot \boldsymbol{\rho}]. \tag{24}$$

Esta aberração resulta em um deslocamento da imagem, mas não causa degradação da imagem.

Conforme pode ser observado, ambas as teorias apresentadas resultam em conclusões similares a respeito dos efeitos de perturbações em sistemas ópticos. A teoria de Thompson, apresentada primeiro em 1980, abrange descentralização e inclinação e apresenta resultados inéditos a respeito dos campos de aberrações nestes sistemas. Pesquisas recentes encontradas na literatura mostram que essa teoria consegue prever e explicar satisfatoriamente os efeitos de erros de alinhamento em sistemas variados (20, 23, 41, 42).

A teoria de Sasián visa apresentar esses efeitos de uma forma mais simples e intuitiva, quando eles ocorrem em sistemas desalinhados ou sem simetria axial, mas que ainda preserva alguma simetria e se encaixa em sistemas plano-simétricos ou biplano-simétricos.

Ambas foram apresentadas por serem abordagens diferentes e podem ser aplicadas de acordo com o sistema que se pretende analisar.

2.3 Polinômios de Zernike

A seguir, será apresentada uma breve descrição dos polinômios de Zernike. O texto baseia-se nos trabalhos de Virendra N. Mahajan (4, 43), e James C. Wyant e K. Creath (44).

Os polinômios de Zernike foram inicialmente introduzidos por Frits Zernike para o seu método de contraste de fase usado em testes de superfícies de espelhos circulares. O ordenamento dos termos e a sua forma podem aparecer de diversas maneiras, já que não há uma forma padrão. A sua forma ortonormal e o primeiro ordenamento dos termos foram sugeridos por Noll e serão adotados nesta seção (4).

A mesma ordenação é também adotada pelo software de projeto óptico ZEMAX® empregado neste trabalho quando se trata dos polinômios de Zernike definidos como *Zernike Standard Coefficients* pelo software (45).

Polinômios de Zernike representam aberrações otimamente balanceadas, em que os termos de ordens mais altas nas coordenadas da pupila são balanceados com as aberrações de ordens iguais ou mais baixas para minimizar a sua variância ao longo da pupila (43).

O uso dos polinômios ortonormais na expansão de uma função de aberração apresenta a vantagem de que os coeficientes dos termos da expansão representam os seus desvios padrões (4).

Esse polinômio possui propriedades de simetria rotacional podendo ser representado como uma multiplicação de uma função radial ($R(\rho)$) por uma função angular ($G(\theta)$). A função angular é contínua, repete-se a cada 2π radianos e satisfaz o requisito de que, quando as coordenadas do sistema são rotacionadas por um determinado ângulo, a forma polinomial não é alterada. A função radial deve ser um polinômio em ρ de grau *n* cuja potência não pode ser menor que *m*. Ainda, $R(\rho)$ é par, se *m* é par e ímpar, se *m* é ímpar (44).

A equação (25) apresenta a forma ortonormal de um termo da aberração de onda $W(\rho,\theta)$ expandida em termos dos polinômios de Zernike:

$$W_n^m(\rho,\theta;\epsilon;\gamma) = c_{nm}\left[\frac{2(n+1)}{1+\delta_{m0}}\right]^{\frac{1}{2}} R_n^m(\rho;\epsilon;\gamma)\cos m\theta, \qquad (25)$$

onde

- c_{nm} é o coeficiente de aberração e representa o desvio padrão da aberração (a menos que $n=m=0$;

- $R_n^m(\rho;\epsilon;\gamma)$ representa um polinômio de Zernike radial;

- *n* e *m* são inteiros positivos incluindo zero, $n-m\geq 0$ e par;

- (r,θ) são as coordenadas polares de um ponto no plano da pupila, $0\leq\theta<2\pi$;

- ϵ é a razão de obscuração linear na pupila, $\epsilon=0$ corresponde a uma pupila circular;

- $\rho = r/a$ é a coordenada radial normalizada, $\epsilon\leq\rho\leq 1$;

- ϵa e *a* são os raios interno e externo de uma pupila anular;

- $\gamma = (a/\omega)^2$ é o razão de truncamento de uma pupila gaussiana, $\gamma = 0$ corresponde a uma pupila uniformemente iluminada;

- ω é a distância radial conhecida como raio gaussiano, em que a irradiância da pupila cai a e^{-2} de seu valor no centro na ausência de obscuração;

- δ_{ij} é o delta de Kronecker.

A ortonormalidade dos polinômios de Zernike implica que:

$$\int_{\epsilon}^{1}\int_{0}^{2\pi} W_n^m(\rho,\theta;\epsilon;\gamma) W_{n'}^{m'}(\rho,\theta;\epsilon;\gamma) A(\rho)\rho d\rho d\theta \Big/ \int_{\epsilon}^{1}\int_{0}^{2\pi} A(\rho)\rho d\rho d\theta = c_{nm}^2 \delta_{nn'}\delta_{mm'} \qquad (26)$$

e a distribuição da amplitude da radiação que ilumina a pupila $A(\rho)$ é dada por:

$$A(\rho) = A_0 \exp(-\gamma\rho^2), \qquad (27)$$

onde A_0 é uma constante e $[A_0^2] = W/m^2$ (unidade de irradiância).

No caso de sistemas ópticos com pupilas circulares de raio a $(0 \le \rho \le 1, 0 \le \theta < 2\pi)$, $\epsilon = 0$, $\gamma = 0$, o conjunto completo dos polinômios circulares de Zernike é dado por:

$$W(\rho,\theta) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\left[\frac{2(n+1)}{1+\delta_{m0}}\right]^{\frac{1}{2}} R_n^m(\rho).(c_{nm}\cos m\theta + s_{nm}\operatorname{sen} m\theta), \qquad (28)$$

onde c_{nm} e s_{nm} são os coeficientes de aberração. O polinômio radial é dado por:

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\left(\frac{n+m}{2}-s\right)!\left(\frac{n-m}{2}-s\right)!}\rho^{n-2s}. \qquad (29)$$

O polinômio radial é par ou ímpar em ρ, dependendo de n (ou m) ser par ou ímpar, respectivamente. O polinômio de grau n em ρ contém os termos $\rho^n, \rho^{n-2}, \ldots, \rho^m$, e:

$$R_n^m(0) = \begin{cases} \delta_{m0}, & \frac{n}{2} \text{ par} \\ -\delta_{m0}, & \frac{n}{2} \text{ ímpar} \end{cases}, \qquad R_n^m(1) = 1, \qquad R_n^n(\rho) = \rho^n. \qquad (30)$$

O índice n representa a ordem ou o grau do polinômio; já que determina a maior potência de ρ, m pode ser chamado de frequência azimutal.

A ortogonalidade dos polinômios radiais e das funções angulares para sistemas com pupila circular é dada por:

$$\int_0^1 R_n^m(\rho) R_{n'}^m(\rho) \rho d\rho = \frac{1}{2(n+1)} \delta_{nn'}, \tag{31}$$

$$\int_0^{2\pi} \cos m\theta \cos m'\theta d\theta = \pi(1+\delta_{m0})\delta_{mm'}, \tag{32}$$

$$\int_0^{2\pi} \operatorname{sen} m\theta \operatorname{sen} m'\theta d\theta = \pi\delta_{mm'}, \tag{33}$$

$$\int_0^{2\pi} \cos m\theta \operatorname{sen} m'\theta d\theta = 0. \tag{34}$$

Os coeficientes de aberração são dados por:

$$(c_{nm}, s_{nm}) = \left(\frac{1}{\pi}\right)\left[\frac{2(n+1)}{1+\delta_{m0}}\right]^{\frac{1}{2}} \int_0^1 \int_0^{2\pi} W(\rho,\theta) R_n^m(\rho)(\cos m\theta, \operatorname{sen} m\theta)\rho d\rho d\theta. \tag{35}$$

O número de termos de aberração na expansão da função de aberração em termos dos polinômios de Zernike para um dado n é:

$$N_n = (n+1)(n+2)/2. \tag{36}$$

Os termos de aberração de uma expansão em série de potências são conhecidos como aberrações clássicas. Devido a sua ortogonalidade, os termos de aberração de uma expansão em polinômios de Zernike são chamados de aberrações ortogonais ou aberrações balanceadas (4).

O termo *piston* ($n = m = 0$) não é considerado uma aberração, mas está incluido no número total de termos obtido pela equação (36).

As aberrações ortogonais são constituídas de uma ou mais aberrações clássicas.

Os polinômios de Zernike são únicos no sentido de que eles formam o único conjunto completo de polinômios em duas coordenadas (ρ, θ), que são ortogonais

em um círculo unitário; são invariantes em forma com relação à rotação dos eixos em torno da origem e apresentam um polinômio para cada par possível de n e m.

Esses polinômios são amplamente utilizados em projetos e testes ópticos, por expressarem uma função de aberração por causa da associação com as aberrações clássicas otimamente balanceadas.

A função de aberração em termos dos Polinômios Circulares de Zernike também pode ser escrita como:

$$W(\rho,\theta)=\sum_{j=1}^{\infty}a_j Z_j(\rho,\theta), \tag{37}$$

onde

$$Z_{j\,par}(\rho,\theta)=\sqrt{2(n+1)}R_n^m(\rho)\cos m\theta, \qquad m\neq 0, \tag{38}$$

$$Z_{j\,impar}(\rho,\theta)=\sqrt{2(n+1)}R_n^m(\rho)\operatorname{sen} m\theta, \qquad m\neq 0 \tag{39}$$

$$Z_j(\rho,\theta)=\sqrt{n+1}R_n^0(\rho), \qquad m=0. \tag{40}$$

Nessas equações, $_j$ é o número de ordenamento polinomial, é função de n e m; a_j é o coeficiente de aberração ou de expansão[f]. O ordenamento dos termos não implica que os coeficientes de aberração diminuam quando o valor de n aumenta. Os termos com j par representam polinômios simétricos ($\cos m\theta$), e os termos com j ímpar representam polinômios antissimétricos ($\operatorname{sen} m\theta$) e, para um dado n, polinômios com valores menores de m são ordenados primeiro.

A ortonormalidade dos polinômios de Zernike implica que:

$$\int_0^1\int_0^{2\pi}Z_j(\rho,\theta)Z_{j'}(\rho,\theta)\rho\,d\rho\,d\theta \Big/ \int_0^1\int_0^{2\pi}\rho\,d\rho\,d\theta=\delta_{jj'}. \tag{41}$$

O coeficiente a_j é dado por:

$$a_j=\pi^{-1}\int_0^1\int_0^{2\pi}W(\rho,\theta)Z_j(\rho,\theta)\rho\,d\rho\,d\theta. \tag{42}$$

[f] Esta é a notação usada por Mahajan (4). Wyant (44) usa Z_j para o coeficiente do termo de número j. A notação de Wyant será adotada adiante nesta tese.

A Tabela 4 exibe os 37 primeiros termos dos polinômios de Zernike ortonormais. A Figura 5 apresenta gráficos tridimensionais de alguns termos dos polinômios de Zernike; a Figura 6 apresenta gráficos de densidade destes mesmos termos (44).

Para o caso de um sistema rotacionalmente simétrico em torno do eixo óptico, a função de aberração também é simétrica em torno do plano tangencial (plano que contém o eixo óptico e o objeto pontual para o qual as aberrações estão sendo calculadas). Se o ângulo θ é medido a partir do plano tangencial, $\text{sen}m\theta = 0$; assim, os coeficientes S_{nm} são nulos. Dessa forma, em sistemas rotacionalmente simétricos apenas os termos em $\cos m\theta$ precisam ser considerados (4). Neste caso, o número de termos é dado por:

$$N_n = (n+2)(n+4)/8. \tag{43}$$

No caso dos termos de quarta ordem, $n = 4$, são obtidos seis termos: o *piston* e as cinco aberrações primárias clássicas.

Erros de fabricação e distorções térmicas em sistemas rotacionalmente simétricos introduzem aberrações que consistem em ambos os termos $\cos m\theta$ e $\text{sen}m\theta$. Além disso, alguns erros introduzidos no processo de fabricação não podem ser bem representados utilizando-se poucos termos destes polinômios (44).

O valor médio e o valor quadrático médio da função de aberração são dados respectivamente por (4):

$$\left\langle W(\rho,\theta)\right\rangle = \int_0^1\int_0^{2\pi} W(\rho,\theta)\,\rho d\rho d\theta \Big/ \int_0^1\int_0^{2\pi} \rho d\rho d\theta = c_{00}, \tag{44}$$

$$\left\langle W^2(\rho,\theta)\right\rangle = \int_0^1\int_0^{2\pi} W^2(\rho,\theta)\,\rho d\rho d\theta \Big/ \int_0^1\int_0^{2\pi} \rho d\rho d\theta = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\left(c_{nm}^2 + s_{nm}^2\right). \tag{45}$$

Sendo σ_W o desvio padrão, a variância é dada por:

$$\sigma_w^2 = \left\langle W^2(\rho,\theta)\right\rangle - \left\langle W(\rho,\theta)\right\rangle^2 = \sum_{n=1}^{\infty}\sum_{m=0}^{n}(c_{nm}^2 + s_{nm}^2). \tag{46}$$

O valor quadrático médio (RMS) da função de onda é dado por:

$$rms = \left\langle W^2(\rho,\theta)\right\rangle^{1/2}. \tag{47}$$

Este valor é diferente do desvio padrão, se c_{00} é diferente de zero.

Sendo o valor médio de um termo de aberração igual a zero, com exceção do *piston*, o valor RMS é igual ao seu desvio padrão, sendo, portanto, igual ao valor do seu coeficiente de aberração. Desta forma, pode-se escrever:

$$\left\langle W\left(\rho,\theta\right)\right\rangle = a_1, \qquad \left\langle W^2\left(\rho,\theta\right)\right\rangle = \sum_{j=1}^{\infty} a_j^2, \qquad \sigma_w^2 = \sum_{j=2}^{\infty} a_j^2. \tag{48}$$

Conforme mencionado anteriormente, os polinômios de Zernike representam aberrações otimamente balanceadas, em que os termos de ordens mais altas nas coordenadas da pupila são balanceados com as aberrações de ordens iguais ou mais baixas para minimizar a sua variância ao longo da pupila. Desta forma, o termo número 1 é constante e conhecido como *piston*; os termos 2 e 3 representam inclinação; o termo 4 representa *defocus*; 5 e 6 representam astigmatismo e *defocus*; 7 e 8 representam coma e inclinação; o termo 11 representa aberração esférica balanceada com *defocus* (43, 46).

Outra forma dos polinômios de Zernike, também encontrada no programa de projeto óptico adotado (ZEMAX®), é conhecida como *Zernike Fringe Coefficients*, e é também referenciada como *University of Arizona polynomials* por ter sido desenvolvida na Universidade do Arizona (45, 47) (Tabela 5). Possui apenas 37 termos, que não são ortonormais, porém eles são todos normalizados para que tenham um valor unitário na borda da pupila. O ângulo é medido no sentido anti-horário a partir do eixo +*x*. A coordenada radial é normalizada e não tem dimensão (45).

Tabela 4 – Polinômios circulares de Zernike ($x = \rho\cos\theta, y = \rho\operatorname{sen}\theta, 0 \le \rho \le 1, 0 \le \theta < 2\pi$).

j	n	m	$Z_j(\rho, \theta)$	Name
1	0	0	1	*Piston*
2	1	1	$2\rho\cos\theta$	Inclinação em x
3	1	1	$2\rho\operatorname{sen}\theta$	Inclinação y
4	2	0	$\sqrt{3}(2\rho^2 - 1)$	*Defocus*
5	2	2	$\sqrt{6}\rho^2\operatorname{sen}2\theta$	
6	2	2	$\sqrt{6}\rho^2\cos 2\theta$	Astigmatismo
7	3	1	$\sqrt{8}(3\rho^3 - 2\rho)\operatorname{sen}\theta$	Coma primário em y
8	3	1	$\sqrt{8}(3\rho^3 - 2\rho)\cos\theta$	Coma primário em x
9	3	3	$\sqrt{8}\rho^3\operatorname{sen}3\theta$	
10	3	3	$\sqrt{8}\rho^3\cos 3\theta$	
11	4	0	$\sqrt{5}(6\rho^4 - 6\rho^2 + 1)$	Aberração esférica primária
12	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2)\cos 2\theta$	Astigmatismo secundário
13	4	2	$\sqrt{10}(4\rho^4 - 3\rho^2)\operatorname{sen}2\theta$	
14	4	4	$\sqrt{10}\rho^4\cos 4\theta$	
15	4	4	$\sqrt{10}\rho^4\operatorname{sen}4\theta$	
16	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho)\cos\theta$	Coma secundário em x
17	5	1	$\sqrt{12}(10\rho^5 - 12\rho^3 + 3\rho)\operatorname{sen}\theta$	Coma secundário em y
18	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3)\cos 3\theta$	
19	5	3	$\sqrt{12}(5\rho^5 - 4\rho^3)\operatorname{sen}3\theta$	
20	5	5	$\sqrt{12}\rho^5\cos 5\theta$	
21	5	5	$\sqrt{12}\rho^5\operatorname{sen}5\theta$	
22	6	0	$\sqrt{7}(20\rho^6 - 30\rho^4 + 12\rho^2 - 1)$	Aberração esférica secundária
23	6	2	$\sqrt{14}(15\rho^6 - 20\rho^4 + 6\rho^2)\operatorname{sen}2\theta$	
24	6	2	$\sqrt{14}(15\rho^6 - 20\rho^4 + 6\rho^2)\cos 2\theta$	
25	6	4	$\sqrt{14}(6\rho^6 - 5\rho^4)\operatorname{sen}4\theta$	
26	6	4	$\sqrt{14}(6\rho^6 - 5\rho^4)\cos 4\theta$	
27	6	6	$\sqrt{14}\rho^6\operatorname{sen}6\theta$	
28	6	6	$\sqrt{14}\rho^6\cos 6\theta$	
29	7	1	$4(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho)\operatorname{sen}\theta$	Coma terciário em y
30	7	1	$4(35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho)\cos\theta$	Coma terciário em x
31	7	3	$4(21\rho^7 - 30\rho^5 + 10\rho^3)\operatorname{sen}3\theta$	
32	7	3	$4(21\rho^7 - 30\rho^5 + 10\rho^3)\cos 3\theta$	
33	7	5	$4(7\rho^7 - 6\rho^5)\operatorname{sen}5\theta$	
34	7	5	$4(7\rho^7 - 6\rho^5)\cos 5\theta$	
35	7	7	$4\rho^7\operatorname{sen}7\theta$	
36	7	7	$4\rho^7\cos 7\theta$	
37	8	0	$3(70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1)$	Aberração esférica terciária

Fonte: Adaptada de MAHAJAN (4).

Tabela 5 – Polinômios *Fringe* de Zernike[g].

Termo	Z(ρ, φ)	Termo	Z(ρ, φ)	Termo	Z(ρ, φ)
1	1	14	$(10\rho^4 - 12\rho^2 + 3)\rho\cos\varphi$	27	$\rho^5 \operatorname{sen} 5\varphi$
2	$\rho\cos\varphi$	15	$(10\rho^4 - 12\rho^2 + 3)\rho\operatorname{sen}\varphi$	28	$(6\rho^2 - 5)\rho^4 \cos 4\varphi$
3	$\rho\operatorname{sen}\varphi$	16	$20\rho^6 - 30\rho^4 + 12\rho^2 - 1$	29	$(6\rho^2 - 5)\rho^4 \operatorname{sen} 4\varphi$
4	$2\rho^2 - 1$	17	$\rho^4 \cos 4\varphi$	30	$(21\rho^4 - 30\rho^2 + 10)\rho^3 \cos 3\varphi$
5	$\rho^2 \cos 2\varphi$	18	$\rho^4 \operatorname{sen} 4\varphi$	31	$(21\rho^4 - 30\rho^2 + 10)\rho^3 \operatorname{sen} 3\varphi$
6	$\rho^2 \operatorname{sen} 2\varphi$	19	$(5\rho^2 - 4)\rho^3 \cos 3\varphi$	32	$(56\rho^6 - 105\rho^4 + 60\rho^2 - 10)\rho^2 \cos 2\varphi$
7	$(3\rho^2 - 2)\rho\cos\varphi$	20	$(5\rho^2 - 4)\rho^3 \operatorname{sen} 3\varphi$	33	$(56\rho^6 - 105\rho^4 + 60\rho^2 - 10)\rho^2 \operatorname{sen} 2\varphi$
8	$(3\rho^2 - 2)\rho\operatorname{sen}\varphi$	21	$(15\rho^4 - 20\rho^2 + 6)\rho^2 \cos 2\varphi$	34	$(126\rho^8 - 280\rho^6 + 210\rho^4 - 60\rho^2 + 5)\rho\cos\varphi$
9	$6\rho^4 - 6\rho^2 + 1$	22	$(15\rho^4 - 20\rho^2 + 6)\rho^2 \operatorname{sen} 2\varphi$	35	$(126\rho^8 - 280\rho^6 + 210\rho^4 - 60\rho^2 + 5)\rho\operatorname{sen}\varphi$
10	$\rho^3 \cos 3\varphi$	23	$(35\rho^6 - 60\rho^4 + 30\rho^2 - 4)\rho\cos\varphi$	36	$252\rho^{10} - 630\rho^8 + 560\rho^6 - 210\rho^4 + 30\rho^2 - 1$
11	$\rho^3 \operatorname{sen} 3\varphi$	24	$(35\rho^6 - 60\rho^4 + 30\rho^2 - 4)\rho\operatorname{sen}\varphi$	37	$924\rho^{12} - 2772\rho^{10} + 3150\rho^8 - 1680\rho^6 + 420\rho^4 - 42\rho^2 + 1$
12	$(4\rho^2 - 3)\rho^2 \cos 2\varphi$	25	$70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1$		
13	$(4\rho^2 - 3)\rho^2 \operatorname{sen} 2\varphi$	26	$\rho^5 \cos 5\varphi$		

Fonte: Adaptada de RADIANT ZEMAX (45).

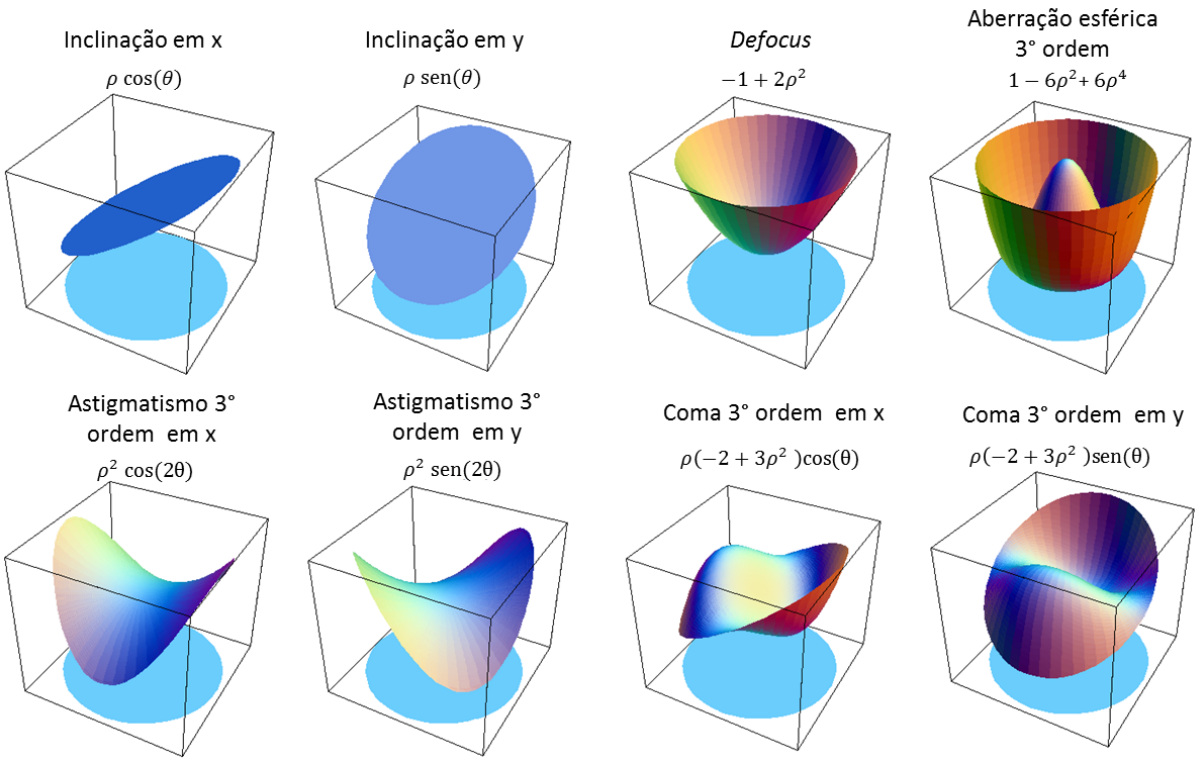

Figura 5 – Polinômios de Zernike. Fonte: Adaptada de WYANT (49).

[g] No que diz respeito ao ângulo, há uma diferença de notação nesta tabela extraída do software Zemax em relação às notações adotadas por Wyant e por Mahajan. Como ela foi extraída do manual do software, que apresenta mais informações a respeito destes polinômios, a notação usada por ele foi mantida, apesar de terem sido feitas algumas modificações para a sua apresentação nesta tese.

Este formato (ordenação) dos polinômios de Zernike é muito utilizado em softwares de interferômetros. O interferômetro Zygo GPI/XP também adota *Zernike Fringe Coefficients* para a ordenação de seus termos, iniciando a numeração dos termos a partir de 0 (48). O software ZEMAX inicia a numeração dos termos a partir do termo de número 1.

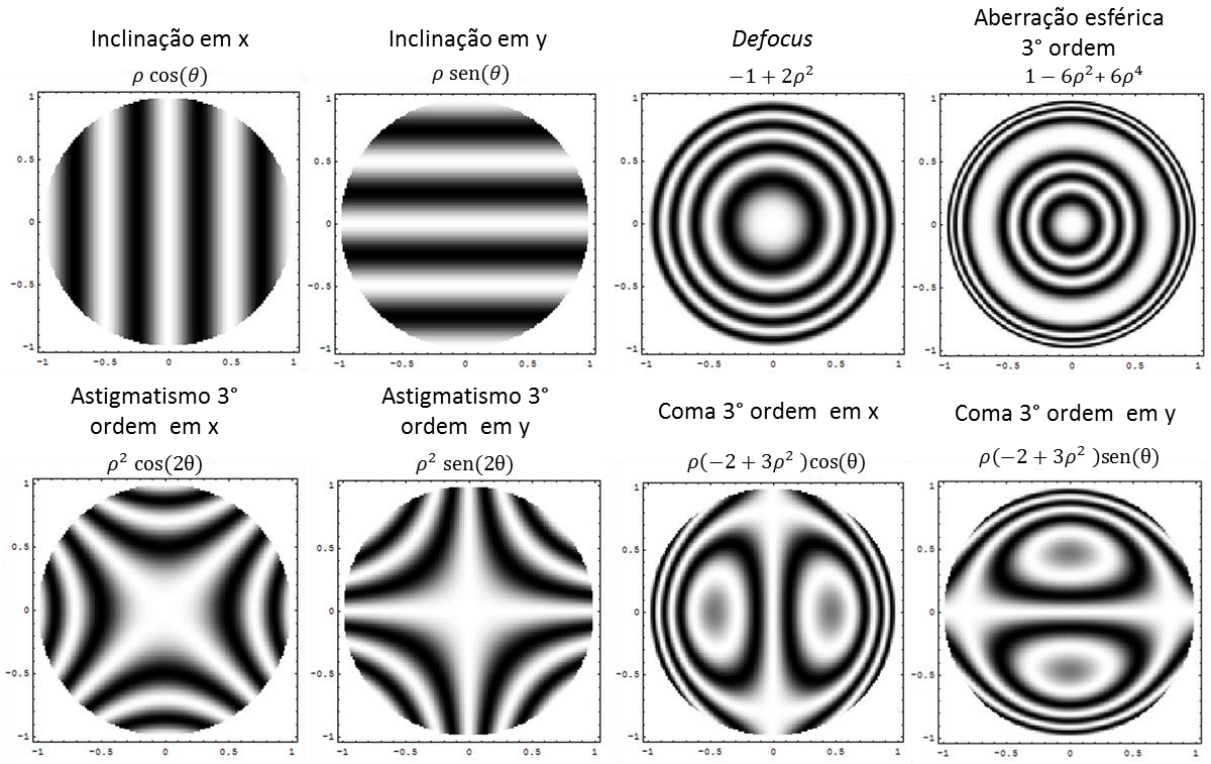

Figura 6 – Gráficos de densidade dos polinômios de Zernike. Fonte: Adaptada de WYANT (50).

O fato de os polinômios de Zernike não apresentarem dependência com o campo faz com que medidas de frente de onda usando interferômetros forneçam dados de aberração em apenas um ponto no campo. Desta forma, a aberração curvatura de campo parecerá *defocus,* e distorção parecerá inclinação. Por esse motivo, devem ser realizadas medidas em diferentes pontos ao longo do campo visando à obtenção efetiva das aberrações do sistema (44).

Tessieres (34), em sua dissertação de mestrado, propõe uma versão modificada de equações que descrevem as aberrações com dependência com o campo, com base na equação de aberração de onda de Thompson, em termos dos polinômios de Zernike com dependência com o campo. Contudo, a proposta de Tessieres de

reescrever os termos de aberração em função desses polinômios levou a termos de aberração que não são ortogonais em todo o campo (6).

Manuel (6), baseada no trabalho prévio de Tessieres, propôs em sua tese de doutorado uma nova forma para as aberrações em função dos polinômios de Zernike, agora ortogonal no campo e na pupila. Sua proposta baseia-se em uma expansão em polinômios de Zernike duplos usando produto de polinômios de Zernike para o campo e polinômios de Zernike para a pupila.

Gray et al. (51) também apresentam expressões para a expansão da função de aberração de onda em termos de polinômios de Zernike com dependência com o campo.

2.4 Defeitos nas superfícies de elementos ópticos

Apesar de as superfícies ópticas serem consideradas perfeitas em grande parte das análises teóricas realizadas em sistemas ópticos, o processo de fabricação e os métodos metrológicos ocasionam imperfeições nessas superfícies (52).

Nos processos convencionais de fabricação, a superfície óptica é gerada e polida. Nesses processos, erros de superfície são gerados podendo apresentar frequências baixas e altas.

Devido à demanda de redução de custo e tempo de fabricação, novas técnicas vêm sendo aprimoradas como, por exemplo, processos de usinagem como o chamado torneamento com ferramenta de diamante de aresta única (*single-point diamond turning*) (53) ou microrretificação determinística *(deterministic microgrinding)* (54). As superfícies ópticas fabricadas com essas técnicas apresentam adicionalmente erros de superfície periódicos de frequências intermediárias. Esses erros de frequências médias e altas ocasionam efeitos de espalhamento que levam à degradação da imagem gerada pelo sistema óptico (54).

O efeito causado por estas imperfeições nas superfícies dos elementos é a degradação do desempenho do sistema óptico. Métodos para descrever esta degradação foram propostos na literatura (54-57).

Apesar de não haver um padrão bem estabelecido para classificação destas imperfeições, diversos autores adotam a seguinte separação dos tipos de erros superficiais (52, 55, 57-58):

- erros de forma ou *figure errors*: são erros topográficos que ocorrem geralmente durante o processo de fabricação e apresentam baixas frequências. Estes erros podem ser tratados por meio das aberrações ópticas convencionais. Podem ser medidos por interferometria e expressos por meio dos polinômios de Zernike (55). No caso de sistemas formadores de imagem, erros de forma, com baixas frequências espaciais têm o efeito de transferir a energia do pico central da PSF[h] (*Point Spread Function*) para os picos secundários mais próximos sem ocasionar o alargamento do pico central, apenas reduzindo a sua intensidade sem mudar a resolução (54). Evans et al. (58) propõem um método qualitativo de análise destes erros de superfície visualizando-os por meio do uso dos polinômios de Zernike, a partir dos 37 termos fornecidos por interferômetros comerciais. Os autores propõem análises gráficas dos termos dos polinômios possibilitando extrair informações adicionais que não podem ser diretamente visualizada pela análise topográfica das frentes de onda;

- erros de frequência espacial média ou *Mid-Spatial Frequency errors* (MSF): adicionalmente aos demais erros, ocorrem durante processos de fabricação com técnicas mais modernas (como microrretificação determinística). Ocasionam ondulações superficiais periódicas regulares residuais chamadas de *ripple* ou MSF (52). Erros MSF apresentam frequências intermediárias entre as frequências que abrangem os erros de forma e de acabamento. Estes erros degradam os picos e nodos da PSF. Estas alterações da PSF ocasionam degradação da qualidade de imagem (52). Erros de médias e altas frequências espalham a luz. Erros MSF levam a espalhamentos de pequenos ângulos (*small-angle scatter*), que causam redução da resolução da imagem (54);

- erros de acabamento ou *finish errors*: nos métodos de fabricação convencionais podem ocorrer fraturas mecânicas de superfície que levam a ondulações de comprimentos de escala espacial curta, chamados de microrrugosidades

[h] PSF em um sistema óptico é a distribuição da intensidade de energia em torno de imagem de uma fonte pontual. Fonte: The Photonics Dictionary, Pittsfield: Laurin Publishing, p. D109, 2004. Maiores informações podem ser encontradas em Smith (17).

(*microroughness*) ou espalhamento (*scatter*). Em sistemas que formam imagem, estes erros com altas frequências causam espalhamento de luz com grandes ângulos (*wide-angle scatter*) que produz o chamado *veiling glare* que reduz a relação sinal ruído e o contraste da imagem e degradam a transmitância do sistema (52, 54).

A Figura 7 apresentada por Aikens, DeGroote e Youngworth (52) mostra uma ilustração (com frequências espaciais arbitrariamente definidas pelos autores) das três regiões de erros com frequências espaciais diferentes. Os autores também propõem uma forma de distinção entre esses erros em função do comprimento de Fresnel (*Fresnel length*). Mais informações podem ser encontradas na referência citada.

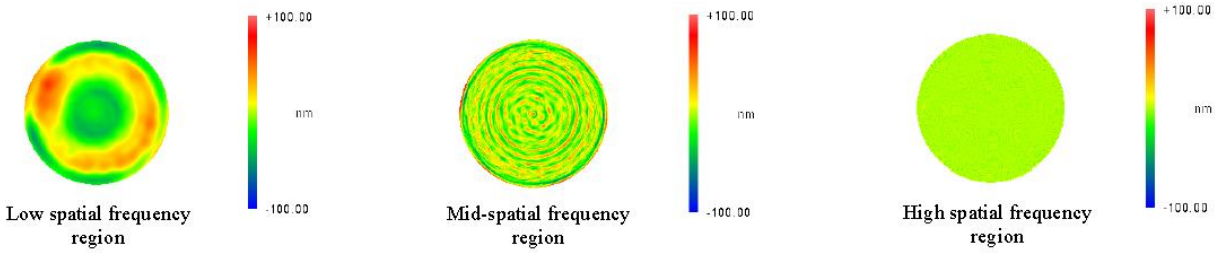

Figura 7 – Superfície fabricada por método de polimento determinístico analisada em três diferentes regiões de frequências espaciais (arbitrariamente escolhidas pelos autores). Fonte: AIKENS, DEGROOTE e YOUNGWORTH (52).

Kotha A. e Harvey J. E. (54) e Harvey J. E., Lewotsky K. L. e Kotha A. (57) apresentaram uma ilustração gráfica dos domínios das frequências espaciais dos três diferentes tipos de erros superficiais (Figura 8). O gráfico mostra a função PSD (*Power spectral density function*) em função da frequência espacial dos erros superficiais. A PSD fornece informações a respeito das frequências espaciais que produzem espalhamento de luz (54). Para cada domínio de frequências, é apresentada uma ilustração dos efeitos causados por estes erros na PSF do sistema.

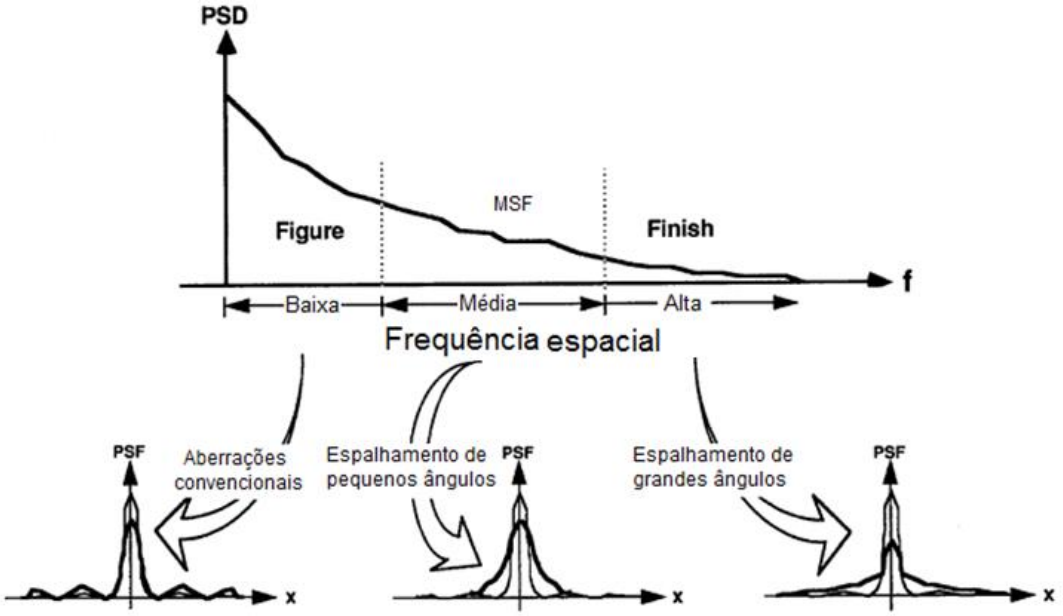

Figura 8 – Domínio dos diferentes tipos de erros superficiais em relação à frequência espacial. Na figura, PSD significa *Power Spectral Density function* ou função de densidade espectral. Fonte: Adaptada de KOTHA e HARVEY (54) e HARVEY, LEWOTSKY, KOTHA (57).

Nos casos apresentados nesta tese, apenas erros de forma serão considerados quando irregularidades superficiais nos elementos ópticos forem incluídas nas análises.

Erros de frequências mais altas não serão abordados pois não interferem diretamente nas aberrações do sistema, podendo ser considerados perdas em sistemas que formam imagem. Da mesma forma erros de frequência espacial média também não serão considerados aqui, apesar de causar degradação no desempenho.

2.4.1 Erros de forma

Erros de superfície podem ser gerados por outros motivos além do processo de fabricação, dentre eles temperatura, pressão, dispositivos e suportes mecânicos, gravidade etc (6, 59).

Em sua tese de doutorado, Manuel (6) apresenta um estudo a respeito dos tipos de erros de forma e sua relação com as demais aberrações de um sistema com dois

espelhos. Além disso, descreve quais erros de forma podem ser compensados, quer seja por erros de forma em outras superfícies, ou por inclinação/descentralização de outros elementos. A seguir será apresentada uma breve descrição das análises realizadas por ela baseadas no trabalho prévio de Tessieres.

Em um telescópio com dois espelhos, o espelho primário é o maior e é nele que geralmente está posicionada a pupila de entrada do sistema. Erros de forma nesse elemento afetam todos os ângulos do campo de forma similar. Se houver um erro de forma do tipo coma ou astigmatismo no espelho primário, haverá coma ou astigmatismo uniforme, respectivamente, em todo o campo.

Como coma uniforme é uma das principais aberrações causadas por desalinhamentos do espelho secundário, esse erro de forma no espelho primário pode ser compensado por meio de um desalinhamento do espelho secundário.

No caso do espelho secundário, ou outros espelhos de um sistema mais complexo, como esse elemento não está localizado em uma pupila do sistema, os efeitos de erros de forma neste elemento não são simples de ser avaliados. No caso do espelho secundário, o feixe de luz com um dado ângulo de entrada, atingirá apenas parte de sua superfície. Por isso, as aberrações no sistema causadas por esses erros de forma, não serão, necessariamente, iguais aos erros de forma presentes na superfície.

Tessieres (34) e Manuel (6) descrevem estes erros em termos dos polinômios de Zernike na forma vetorial, e dos polinômios duplos de Zernike com dependência com o campo.

Dentre as conclusões de seus trabalhos, consta que erros de forma sempre causam erros de foco e/ou astigmatismo.

Por exemplo, para um erro de forma do tipo coma no espelho secundário (ou em um outro elemento que não esteja na pupila do sistema) as aberrações resultantes no sistema são: coma uniforme, astigmatismo linear e inclinação no plano de foco.

A forma funcional do astigmatismo linear gerado por um erro de forma do tipo coma é o mesmo que o gerado por um erro de alinhamento e depende linearmente do desalinhamento. O mesmo ocorre para o coma uniforme e a inclinação do plano focal. Isso indica que esse tipo de erro, quando pequeno, pode ser compensado por um desalinhamento intencional nos elementos.

Por outro lado, ocorrendo erros de forma do tipo *trefoil*[i] em uma superfície que não esteja na pupila, a ocorrência ocasionará *trefoil* uniforme e astigmatismo linear no sistema. *Trefoil* uniforme não é facilmente introduzido por um erro de alinhamento (apesar de ser possível) e o astigmatismo linear causado por esse erro de forma também não pode ser compensado por desalinhamentos, pois possui características diferentes das do astigmatismo linear causado por erros de alinhamento (6).

Observa-se que alguns tipos de erros de forma podem ser compensados por desvios de alinhamento e outros não. Isso indica que erros de forma que ocasionam aberrações similares às de erros de alinhamento podem confundir a RNA e ocasionar erros nas estimativas dos desalinhamentos do sistema.

2.5 Métodos Interferométricos

Testes interferométricos vêm sendo muito usados ao longo dos anos em metrologia óptica. Na área de óptica, interferometria pode ser aplicada como uma ferramenta no processo de fabricação, testes finais e alinhamento de sistemas ópticos (60).

Interferogramas possibilitam a observação das aberrações presentes em um sistema óptico. As diferentes aberrações presentes no sistema geram diferentes padrões de interferência (36) (Figura 9).

Uma forma de se extrair informações dos tipos e magnitudes das aberrações presentes em uma frente de onda é por meio do ajuste dos polinômios de Zernike a esta superfície (61).

Interferômetros utilizam o fenômeno de interferência e funcionam da seguinte forma: um feixe de luz é dividido em duas ou mais partes que percorrem caminhos diferentes; basicamente, um dos feixes é o de referência, e o outro o de medida, que atravessa o sistema que se deseja analisar; os feixes são recombinados e formam

[i] *Trefoil* são erros inseridos por suportes com três pontos (*three point mount-induced error*). Fonte: Fuerschbach, K.; Rolland, J. P.; Thompson, K. P. Extending nodal aberration theory to include mount-induced aberrations with application to freeform surfaces. **Optics Express**, v. 20, n. 18, p. 20139-20155, 2012.

os padrões interferométricos (36, 60). A equação que descreve a interferência entre dois feixes é dada por (61-63):

$$I(x,y) = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\alpha_1 - \alpha_2), \quad (49)$$

onde I é a irradiância (dada pelo quadrado da amplitude do campo eletromagnético), α é a fase da onda e $\alpha_1 - \alpha_2$ é a diferença de fase entre os feixes e $\alpha_1 - \alpha_2 = (\frac{2\pi}{\lambda})(OPD)$ (63).

As propriedades geométricas dos padrões interferométricos são determinadas pela diferença no caminho óptico percorrido pelas duas frentes de onda. Interferômetros medem a diferença de caminho óptico em unidades de comprimento de onda λ da luz utilizada (60).

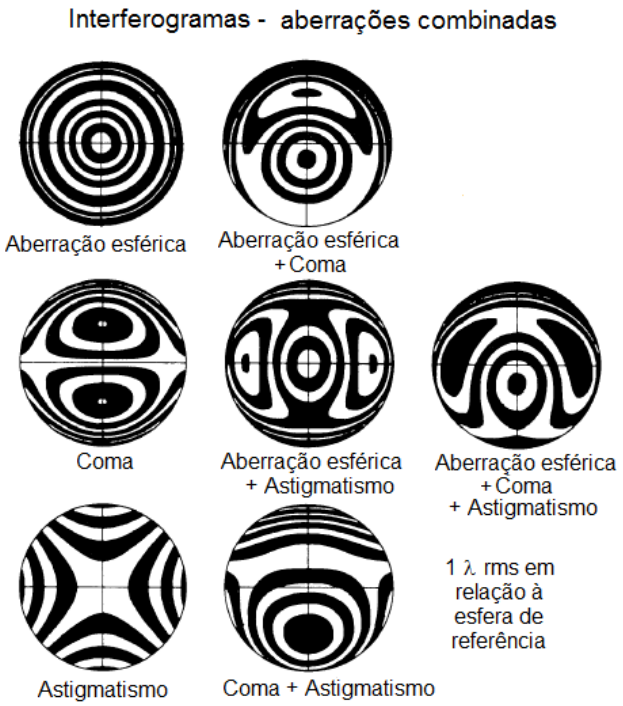

Figura 9 – Interferogramas das aberrações de primeira ordem e de combinações delas. Fonte: Adaptada de ZYGO Corp (60).

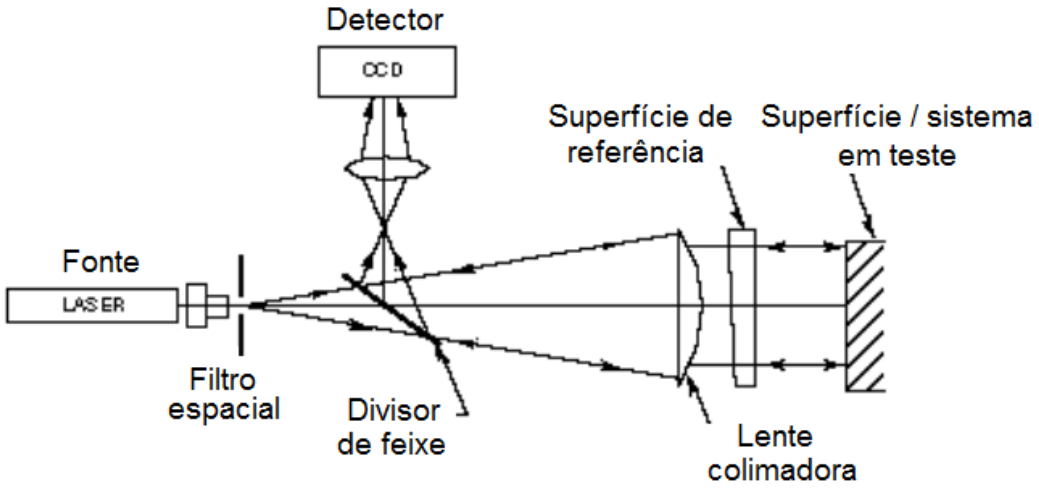

Figura 10 – Esquema ilustrativo de um interferômetro de Fizeau. Fonte: Adaptada de GEARY (61).

Um padrão interferométrico tem um determinado fator de escala, em que um espaçamento entre franjas corresponde a $\lambda/2$, λ ou um múltiplo de λ. O fator de escala depende da configuração de medida.

Existem diversos tipos de interferômetros que podem ser usados para a extração do interferograma de um sistema óptico, um deles é o chamado interferômetro de Fizeau, apresentado esquematicamente na Figura 10.

Maiores informações podem ser encontradas em Mahajan (36), Zygo (60), Geary (61), e Goodwin e Wyant (63).

2.6 Redes Neurais Artificiais - RNA

2.6.1 Introdução e conceitos básicos

Esta seção apresenta uma breve revisão dos principais conceitos relativos a redes neurais artificiais, em especial a um tipo de arquitetura denominada Perceptron. O objetivo é apresentar aqui os princípios básicos que foram usados no desenvolvimento deste trabalho.

As informações apresentadas neste trabalho são baseadas principalmente nos livros de Haykin (15) e de Silva, Spatti e Flauzino (64).

Redes Neurais Artificiais (RNA) enquadram-se, juntamente com Fuzzy, computação evolutiva e diversas outras ferramentas em uma área conhecida como sistemas inteligentes (conexionistas) ou inteligência computacional (64).

Devido a suas características, as RNAs são aplicadas na solução dos mais variados problemas, sendo usadas em diversas áreas do conhecimento. Podem ser empregadas para reconhecimento de padrões, modelagem, processamento de sinais, análise de séries temporais, entre outras (15).

RNAs são modelos computacionais inspirados no sistema nervoso dos seres vivos. São constituídas por um conjunto de unidades de processamento denominados neurônios artificiais, interligados por um grande número de interconexões, as sinapses artificiais. Possuem a habilidade de aprender os comportamentos envolvidos em um sistema a partir de informações extraídas deste e generalizar. A generalização consiste no fato da rede neural produzir saídas adequadas a dados de entrada que não estavam presentes durante o seu processo de aprendizagem (15). Desta forma, RNAs conseguem mapear sistemas não lineares (64).

De forma resumida, algumas das principais características (ou benefícios) das redes neurais artificiais são (15, 64):

- adaptabilidade: a rede é capaz de ajustar os seus parâmetros internos (pesos sinápticos) por meio da apresentação sucessiva de amostras de treinamento ou exemplos da tarefa relacionados ao comportamento do processo, levando a rede à aquisição do conhecimento por experimentação;
- capacidade de aprendizado: a rede neural é capaz de “entender” a relação existente entre as variáveis que compõem o problema. Isso é feito por meio de um processo de treinamento que apresenta à rede os exemplos da tarefa. Cada exemplo é composto por um sinal de entrada único e de uma resposta desejada correspondente. A rede tem a capacidade de aprender dos exemplos apresentados a ela;
- habilidade de generalização: a rede neural, após ser treinada, é capaz de gerar uma saída consistente a um sinal de entrada que não havia sido apresentado a ela durante o processo de treinamento;

- tolerâncias a falhas: uma rede neural consegue ser tolerante a falhas quando parte de sua estrutura interna é corrompida, ou seja, ela exibe uma degradação suave do desempenho, sendo capaz de realizar computação robusta. Assim, devido à natureza distribuída da informação armazenada na rede, é necessário um dano muito extenso para ocasionar uma grande degradação do seu desempenho;
- não linearidade: uma rede constituída por conexões de neurônios não lineares é também não linear. Esta não linearidade é distribuída por toda a rede e é uma propriedade importante, principalmente quando o sistema/problema em que a rede está sendo aplicada possuir comportamento não linear.

A seguir, serão descritos os princípios básicos relacionados às RNAs.

Neurônio artificial

Os neurônios artificiais constituem modelos simplificados inspirados nos neurônios biológicos. O modelo mais simples que apresenta as características de uma rede neural biológica (paralelismo e alta conectividade) foi proposto, em 1943, por McCulloch e Pitts (64).

Os neurônios artificiais são não lineares e fornecem saídas tipicamente contínuas. Eles são capazes de coletar os sinais em suas entradas e produzir uma resposta de acordo com sua função de ativação (64). A Figura 11 apresenta o modelo não linear de um neurônio.

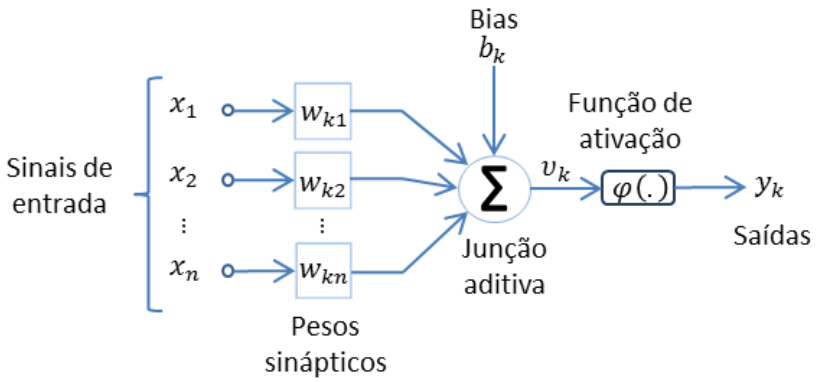

Figura 11 – Modelo de neurônio não linear. Fonte: HAYKIN (15).

O neurônio artificial é constituído de alguns elementos básicos:

- sinais de entrada: os sinais de entrada, dados por $\{x_1, x_2, \ldots, x_n\}$, são obtidos a partir do sistema a que se deseja aplicar a rede neural e consistem em valores assumidos pelas variáveis envolvidas na aplicação específica. Esses dados são normalizados com o objetivo de aumentar a eficiência computacional dos algoritmos de aprendizagem (64);
- sinapses ou elos de conexão: dados por $\{w_{k1}, w_{k2}, \ldots, w_{kn}\}$, constituem os valores usados para ponderar as variáveis de entrada da rede, possibilitando avaliar sua relevância no que diz respeito à funcionalidade do neurônio. Um sinal x_j na sinapse j conectada ao neurônio k é multiplicado por um peso sináptico w_{kj};
- somador ou combinador linear: dado por $\sum$, realiza a soma dos sinais de entrada, ponderados pelos respectivos pesos sinápticos buscando obter um valor de potencial de ativação;
- limiar de ativação ou *bias*: indicado por b_k, especifica qual o patamar apropriado para que o resultado dado pelo combinador linear possa gerar um valor de disparo em relação à saída do neurônio (64). O *bias* ocasiona o aumento ou diminuição da entrada líquida da função de ativação (15);
- potencial de ativação: o potencial de ativação é dado por ($\upsilon_k = u_k + b_k$), onde o resultado do combinador linear é u_k (15). Silva, Spatti e Flauzino descrevem $\upsilon_k = u_k - \theta$, onde θ é o limiar de ativação e, se υ é positivo, o neurônio produz um potencial excitatório, caso contrário, o potencial será inibitório;
- função de ativação: dada por $\varphi(.)$, tem a função de limitar a saída do neurônio dentro de um intervalo de valores dado pela função escolhida;
- sinal de saída: os sinais de saída (y_k) são as saídas produzidas pelo neurônio para um conjunto de sinais de entradas.

As seguintes equações descrevem um neurônio k:

$$u_k = \sum_{j=1}^{n} w_{kj} x_j, \tag{50}$$

$$y_k = \varphi(u_k + b_k). \tag{51}$$

Função de ativação

A função de ativação pode ser de vários tipos, dentre eles:

- função degrau:

  Também conhecida como função de Heaviside, é dada pela função:

$$\varphi(\upsilon) = \begin{cases} 1 & se \quad \upsilon \geq 0 \\ 0 & se \quad \upsilon < 0 \end{cases}. \tag{52}$$

  Para esta função, a saída do neurônio assumirá valor unitário positivo quando o potencial de ativação do neurônio for maior ou igual a zero; caso contrário, assumirá valor nulo.

  Um neurônio com essa função de ativação é conhecido na literatura como o modelo de McCulloch-Pitts (15).

- função linear:

  A função linear (ou identidade) é dada pela equação (53), e os seus valores de saída são dados pelos próprios valores do potencial de ativação:

$$\varphi(\upsilon) = \upsilon. \tag{53}$$

- função logística:

  É dada por:

$$\varphi(\upsilon) = \frac{1}{1 + \exp(-\beta\upsilon)}. \tag{54}$$

  Os resultados de saída da função logística são valores reais entre 0 e 1. O valor da constante β está associado à inclinação da curva.

- função tangente hiperbólica:

  A tangente hiperbólica fornece valores de saída reais no intervalo entre -1 e 1, e da mesma forma que na função logística, a constante β está associada à inclinação da curva. Esta função é dada pela equação:

$$\varphi(\upsilon) = \frac{1 - \exp(-\beta\upsilon)}{1 + \exp(-\beta\upsilon)} \quad . \tag{55}$$

A Figura 12 apresenta as curvas das função citadas anteriormente.

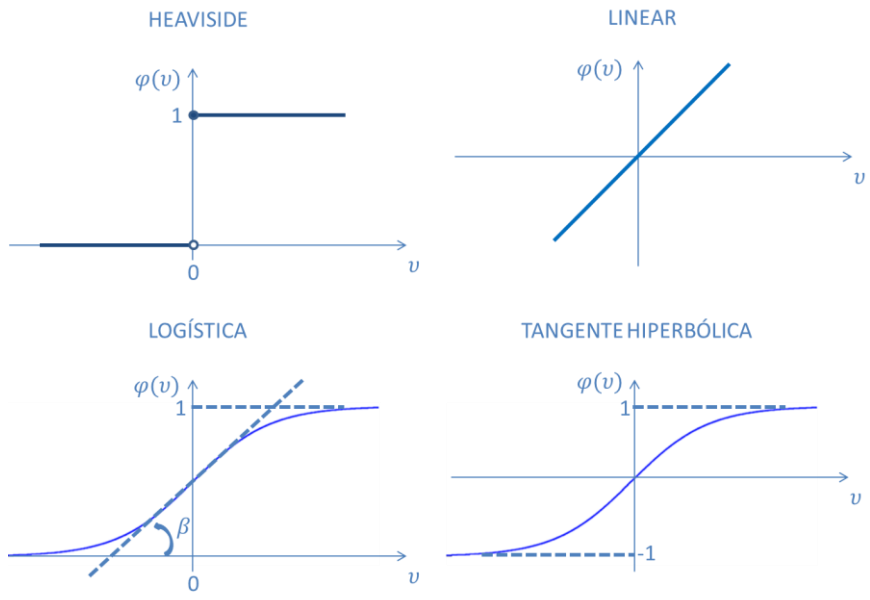

Figura 12 – Representação gráfica das funções de ativação: Heaviside, linear, logística e tangente hiperbólica.

Treinamento supervisionado

Consiste em um tipo de treinamento em que as amostras de treinamento são compostas por pares de entradas e saídas, ou seja, para cada conjunto de dados de entrada $\{x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\}$ apresentado à rede, é também apresentada a saída esperada $\{y_1^{(k)}, y_2^{(k)}, \ldots, y_m^{(k)}\}$. Assim, nesse tipo de treinamento há a necessidade de se apresentar uma tabela de dados (entradas/saídas) representativos do processo, conhecida como tabela atributos/valores.

Assim, o treinamento supervisionado funciona como se houvesse um "professor" ensinando à rede qual a resposta correta para cada amostra de entrada apresentada.

Por meio desse conjunto de treinamento, os pesos sinápticos e limiares são continuamente ajustados. O ajuste é realizado por meio de ações comparativas realizadas pelo algoritmo de aprendizagem. A diferença entre a resposta da rede e a

resposta desejada é utilizada no procedimento de ajuste. A rede é considerada treinada quando a diferença estiver dentro de limites aceitáveis.

Arquitetura de rede neural

Define a maneira como estão arranjados os neurônios que compõem a rede. Tais arranjos são estruturados por meio do direcionamento das conexões sinápticas dos neurônios. As redes neurais podem ser divididas em três partes:

- camada de entrada: recebe as informações (dados de entrada);
- camadas intermediárias: também chamadas de camadas escondidas ou ocultas, contêm os neurônios responsáveis por extrair as características associadas ao processo ou sistema. Nessas camadas é realizado quase todo o processo interno da rede (64);
- camada de saída: é responsável pela produção e apresentação dos resultados da rede provenientes das camadas intermediárias.

As principais arquiteturas de rede podem ser divididas em (64):

- arquitetura *feedforward* de camada simples: constituída de apenas uma camada de entrada e uma única camada que corresponde também à de saída, em que o número de neurônios que formam esta camada define o número de saídas da rede. A informação segue uma direção única da camada de entrada para a de saída. Redes com essa arquitetura são usadas em problemas de classificação de padrões e filtragem linear. A rede Perceptron de camada única e Adaline possuem esse tipo de arquitetura;
- arquitetura *feedforward* de camadas múltiplas: Neste caso, além da camada de entrada e da de saída, redes desse tipo possuem pelo menos uma camada escondida. Também nesse tipo de arquitetura, o número de neurônios da camada de saída define o número de saídas da rede. Esse tipo de rede é empregado na solução de problemas de classificação de padrões, otimização, controle de processos, aproximação de funções, dentre outros. São exemplos de redes com essa arquitetura a rede Perceptron multicamada e Redes de Base Radial (RBF);
- arquitetura recorrente ou realimentada: Neste tipo de arquitetura, as saídas dos neurônios são realimentadas como sinais de entrada para outros

neurônios. Redes desse tipo são usadas em sistemas variantes no tempo, tais como previsões de séries temporais, otimização, controle de processos, etc. Dentre as redes que empregam essa arquitetura estão a rede de Hopfield e a rede Perceptron multicamadas com realimentação;

- arquitetura em estrutura reticulada: redes com estrutura reticulada possuem como principal característica o fato de a disposição espacial dos neurônios estar diretamente relacionada ao processo de ajuste de seus pesos sinápticos e limiares. Redes desse tipo podem ser empregadas em problemas de reconhecimento de padrões, otimização de sistemas, e agrupamento (*clustering*). A rede Kohonen possui este tipo de arquitetura.

Topologia de rede

Para uma determinada arquitetura de rede, a sua topologia representa as diferentes formas de composição estrutural que a rede pode assumir. Por exemplo, o número de neurônios pode ser diferente em duas topologias de uma mesma arquitetura de rede, ou o tipo de função de ativação dos neurônios que constituem cada topologia pode variar (64).

Dentre os diferentes tipos de redes neurais que podem ser aplicados na solução do problema em questão, optou-se por utilizar a rede do tipo Perceptron de múltiplas camadas. Esse tipo de rede, apresentada com mais detalhes na próxima seção, possui ampla utilização nos mais variados tipos de problema e mostrou ser bastante eficiente na solução desse problema específico.

2.6.2 Perceptron

Para a descrição do Perceptron multicamadas, que é a rede que efetivamente foi utilizada neste trabalho, serão apresentados os conceitos básicos envolvidos na rede denominada Perceptron, desde o Perceptron de camada única, rede Adaline e finalmente Perceptron multicamadas.

2.6.2.1 Perceptron de camada única

O Perceptron de camada única consiste na forma mais simples de configuração de uma rede neural artificial e foi introduzido por Rosenblatt em 1958. Esse tipo de rede pode ser empregado para classificação de padrões quando estes são linearmente separáveis, ou seja, padrões que se encontram em lados opostos de um hiperplano. O Perceptron constituído de apenas um neurônio é limitado a realizar classificação de padrões com apenas duas classes (15, 64).

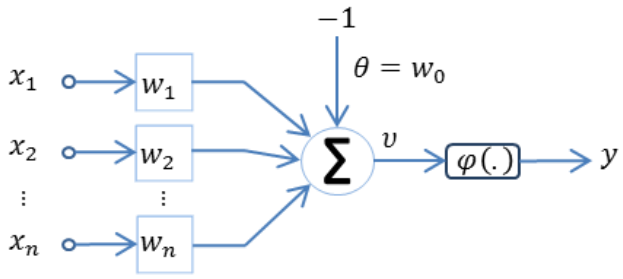

Figura 13 – Ilustração da rede Perceptron. Fonte: SILVA, SPATTI e FLAUZINO (64).

Constituída de apenas uma camada neural contendo apenas um neurônio nesta camada, esta rede se enquadra na arquitetura de rede *feedforward* de camada simples. A Figura 13 apresenta a rede Perceptron contendo n sinais de entrada e uma única saída. A função de ativação da rede Perceptron é a função degrau (que resulta em saídas de valor 0 ou 1) ou degrau bipolar (cujas saídas possíveis são -1 ou 1).

No Perceptron, cada entrada $\{x_i\}$ será inicialmente ponderada pelos pesos sinápticos $\{w_i\}$ de modo que sua importância em relação aos objetivos funcionais do neurônio seja quantificada.

Quando todas as entradas forem ponderadas, o limiar de ativação é adicionado a este conjunto, que consiste no argumento da função de ativação $\varphi(.)$, que, por sua vez, fornece como resultado a saída $\{y\}$ da rede Perceptron. A seguinte formulação matemática é válida (64):

$$\begin{cases} \upsilon = \sum_{i=1}^{n} w_i x_i - \theta, \\ y = \varphi(\upsilon) \end{cases}, \tag{56}$$

onde υ é o potencial de ativação. O *bias* tem o efeito de deslocar o hiperplano ou fronteira de decisão em relação à origem.

O ajuste dos pesos sinápticos e do limiar é realizado por meio do processo de treinamento supervisionado.

No caso do Perceptron, o ajuste dos pesos e limiar é realizado por meio da regra de aprendizado de Hebb.

Aprendizagem hebbiana: Haykin (15) define uma sinapse hebbiana (nome dado em homenagem ao neuropsicólogo Hebb, 1949) como uma sinapse que usa um mecanismo dependente do tempo, altamente local e fortemente interativo, para aumentar a eficiência sináptica como uma função de correlação entre as atividades pré-sináptica e pós-sináptica.

De forma resumida, se a saída do Percepron coincide com a desejada, os pesos e limiares da rede são incrementados proporcionalmente aos valores de seus sinais de entrada. Se a saída da rede é diferente da saída desejada, os pesos e limiares são decrementados. Tal processo é repetido para todas as amostras de treinamento até que a saída do Perceptron seja similar à saída desejada para cada amostra. Em notação algébrica, tem-se:

$$\mathbf{w} \leftarrow \mathbf{w} + \eta \cdot \left(d^{(k)} - y \right) \cdot \mathbf{x}^{(k)}, \tag{57}$$

onde:

$\mathbf{w} = \left[\theta \; w_1 \; w_2 \ldots w_n \right]^T$ é o vetor que contém o limiar e os pesos sinápticos;

$\mathbf{x} = \left[-1 \; x_1^{(k)} \; x_2^{(k)} \ldots x_n^{(k)} \right]^T$ é a *k*-ésima amostra de treinamento;

$d^{(k)}$ é a saída desejada para a *k*-ésima amostra de treinamento;

y é a saída fornecida pelo Perceptron;

η é a taxa de aprendizagem da rede, indica o quão rápido o processo de treinamento da rede estará sendo conduzido rumo à convergência (64).

2.6.2.2 Adaline

Assim como o Perceptron, a rede Adaline é constituída de uma única camada contendo um único neurônio. A composição de várias redes Adaline, formando uma única rede, é denominada Madaline. Esta rede também tem arquitetura do tipo *feedforward* de camada única.

A grande contribuição da rede Adaline, idealizada por Widrow e Hoff em 1960, é a introdução do algoritmo de aprendizado denominado regra Delta.

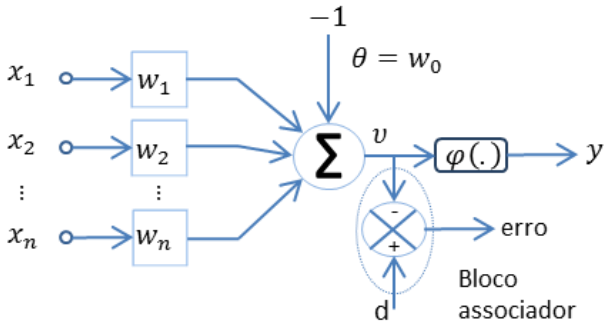

Figura 14 – Ilustração da rede Adaline. Fonte: SILVA, SPATTI e FLAUZINO (64).

Da mesma forma como ocorre na rede Perceptron, os sinais de entrada $\{x_i\}$ são inicialmente ponderados pelos pesos sinápticos $\{w_i\}$, que serão ajustados no processo de treinamento. O potencial de ativação $\{\upsilon\}$ é computado incluindo o limiar, e as saídas $\{y\}$ são produzidas por meio da aplicação da função de ativação, que no caso da rede Adaline é dada pela função degrau ou degrau bipolar. As mesmas funções matemáticas, equação (56), válidas para o Perceptron são também válidas para a rede Adaline; contudo, na rede Adaline há a presença do bloco associador (Figura 14), que tem como função auxiliar no processo de treinamento da rede. Esse bloco gera um sinal de erro que consiste no valor da diferença entre o potencial de ativação produzido pela rede e o valor da saída desejada $\{d\}$, que é usado no ajuste dos pesos sinápticos e limiar da rede.

Nesse tipo de rede, o processo de aprendizado é denominado de regra Delta ou regra de aprendizado de Widrow-Hoff ou, ainda, algoritmo LMS (*least mean square*) ou método do Gradiente Descendente (64).

O objetivo principal da aplicação da regra Delta para o ajuste dos pesos e limiar do neurônio é minimizar a diferença entre a saída desejada $\{d\}$ e a resposta do combinador linear $\{\upsilon\}$, levando em consideração todas as p amostras de treinamento. Para isso, utiliza-se a minimização do erro quadrático $\left\{E\left(\mathbf{w}^*\right)\right\}$ entre υ e d para o ajuste dos pesos $\mathbf{w}=[\theta\ w_1\ w_2 \ldots w_n]^T$ buscando a obtenção de um $\mathbf{w}^*$ ótimo. Assim:

$$E\left(\mathbf{w}^*\right) \leq E\left(\mathbf{w}\right), \quad para \forall \mathbf{w} \in \Re^{n+1}. \tag{58}$$

Define-se a função de erro quadrático em relação às p amostras de treinamento por:

$$E\left(\mathbf{w}\right)=\frac{1}{2}\sum_{k=1}^{p}\left(d^{(k)}-\upsilon\right)^2, \tag{59}$$

ou ainda:

$$E\left(\mathbf{w}\right)=\frac{1}{2}\sum_{k=1}^{p}\left(d^{(k)}-\left(\mathbf{w}^T\cdot\mathbf{x}^{(k)}-\theta\right)\right)^2. \tag{60}$$

Em seguida, para a obtenção do valor ótimo para o erro quadrático, aplica-se o operador gradiente em relação ao vetor $\mathbf{w}$:

$$\nabla E\left(\mathbf{w}\right)=\frac{\partial E(\mathbf{w})}{\partial \mathbf{w}}=\sum_{k=1}^{p}\left(d^{(k)}-\left(\mathbf{w}^T\cdot\mathbf{x}^{(k)}-\theta\right)\right)\cdot\left(-\mathbf{x}^{(k)}\right), \tag{61}$$

assim:

$$\nabla E\left(\mathbf{w}\right)=-\sum_{k=1}^{p}\left(d^{(k)}-\upsilon\right)\cdot\mathbf{x}^{(k)}. \tag{62}$$

Como o objetivo é minimizar o erro quadrático, a atualização dos pesos deve ser realizada na direção oposta à do gradiente; assim, a variação do vetor peso é dada por:

$$\Delta \mathbf{w} = -\eta \cdot \nabla E\left(\mathbf{w}\right) = \eta \sum_{k=1}^{p} \left(d^{(k)} - \upsilon\right) \cdot \mathbf{x}^{(k)} \ . \tag{63}$$

A atualização de $\mathbf{w}$ é dada por:

$$\mathbf{w}^{atual} = \mathbf{w}^{anterior} + \eta \sum_{k=1}^{p} \left(d^{(k)} - \upsilon\right) \cdot \mathbf{x}^{(k)} \ , \tag{64}$$

ou, ainda, em forma discreta e notação algébrica:

$$\mathbf{w} \leftarrow \mathbf{w} + \eta \cdot \left(d^{(k)} - \upsilon\right) \cdot \mathbf{x}^{(k)} \ . \tag{65}$$

Aqui, novamente η é a taxa de aprendizagem que normalmente apresenta valores compreendidos no intervalo $0 < \eta < 1$.

O critério de parada é dado em relação ao erro quadrático médio (EQM) $\left\{E_{qm}(\mathbf{w})\right\}$ em relação a todas as amostras de treinamento:

$$E_{qm}\left(\mathbf{w}\right) = \frac{1}{p} \sum_{k=1}^{p} \left(d^{(k)} - \upsilon\right)^2 . \tag{66}$$

O algoritmo converge quando o erro quadrático médio entre dois passos (épocas) sucessivos for suficientemente pequeno, ou seja, a diferença é menor que a precisão ε requerida:

$$\left| E_{qm}\left(\mathbf{w}^{atual}\right) - E_{qm}\left(\mathbf{w}^{anterior}\right) \right| \leq \varepsilon. \tag{67}$$

O número de épocas de treinamento visa quantificar o número de reapresentações do conjunto de amostras de treinamento para o ajuste do vetor de pesos.

A Figura 15 apresenta uma representação geométrica da regra Delta (64) que mostra os passos de atualização do vetor de pesos rumo ao valor ótimo.

Como pode ser observado na Figura 15, os valores sucessivos de $\mathbf{w}$ ocorrem na direção oposta ao vetor gradiente em relação ao valor anterior de $\mathbf{w}$.

Enquanto no Perceptron o ajuste dos pesos é realizado por meio da regra de Hebb, que considera a resposta produzida após a apresentação individual de cada amostra de treinamento, fazendo com que qualquer hiperplano que esteja posicionado dentro da faixa de separabilidade das classes seja uma solução, no caso da rede Adaline,

que utiliza a regra Delta, que leva em consideração todas as amostras e visa minimizar a diferença entre o valor esperado e o fornecido pela rede, o hiperplano obtido após a convergência da rede será sempre o mesmo. A Figura 16 mostra a representação da diferença entre as soluções da rede Perceptron e da rede Adaline (64).

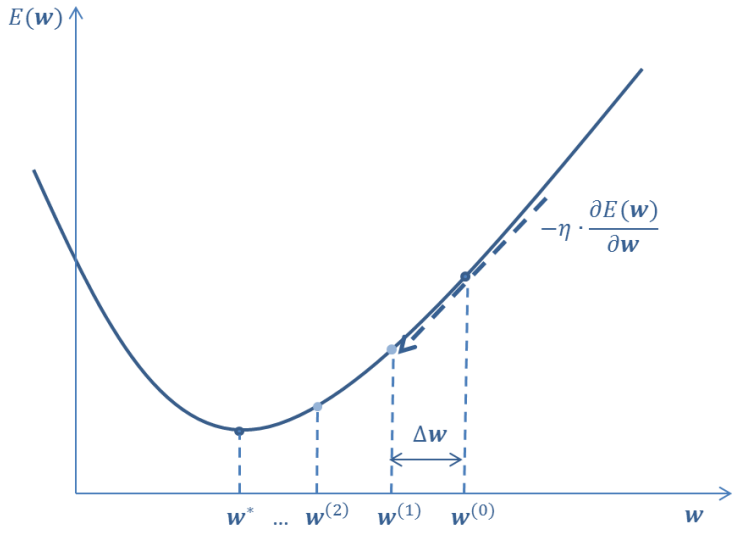

Figura 15 – Interpretação geométrica da regra Delta. Fonte: SILVA, SPATTI e FLAUZINO (64).

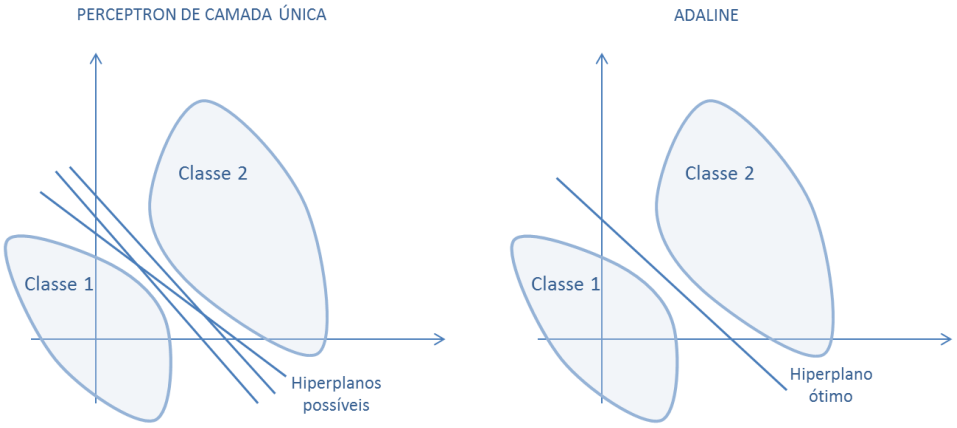

Figura 16 – Diferença entre as soluções das redes Perceptron e Adaline. Fonte: SILVA, SPATTI e FLAUZINO (64).

Desta forma, a rede Adaline é independente dos valores iniciais do seu vetor de pesos; já o Perceptron é fortemente dependente desses valores iniciais, que influenciam a disposição do hiperplano que separa as duas classes envolvidas no problema.

2.6.2.3 Perceptron de múltiplas camadas

A rede Perceptron de múltiplas camadas (PMC ou MLP *multilayer Perceptron*) pertence à arquitetura *feedforward* de camadas múltiplas possuindo, pelo menos, uma camada intermediária. A camada de saída pode ser composta de vários neurônios, cada um deles representando uma das saídas do sistema a ser mapeado. Utiliza treinamento supervisionado, e o algoritmo de aprendizagem é o *back-propagation* (propagação reversa ou algoritmo de retropropagação do erro) também conhecido como regra Delta generalizada.

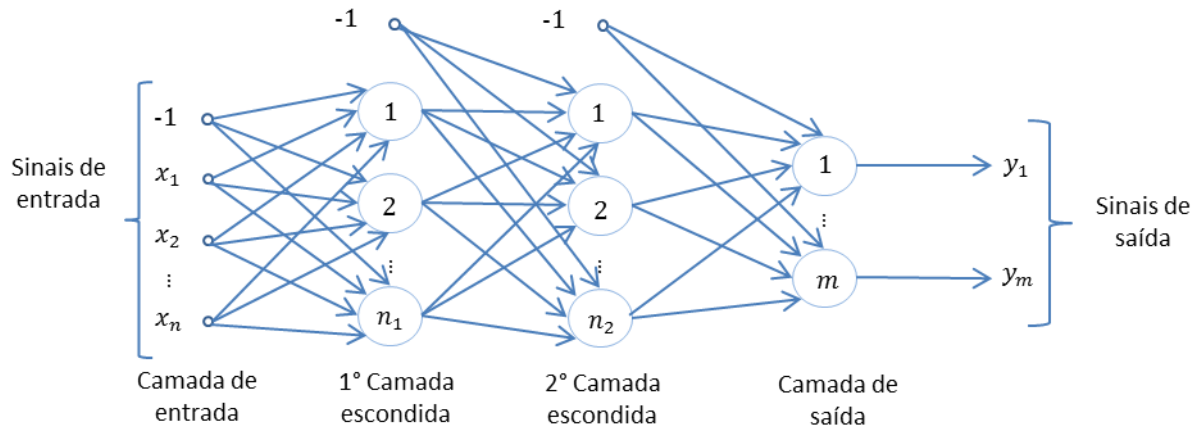

Figura 17 – Ilustração da rede Perceptron multicamadas. Fonte: SILVA, SPATTI e FLAUZINO (64).

A Figura 17 exibe uma ilustração da rede Perceptron multicamadas.

Neste tipo de rede, o conhecimento é distribuído por todos os neurônios da rede. A camada de entrada recebe os sinais/amostras, as camadas intermediárias extraem a maior parte das informações e as codificam por meio dos pesos sinápticos e limiares, e os neurônios da camada de saída recebem os estímulos provenientes da última camada escondida e fornecem as respostas, que serão as saídas da rede.

Assim, três características são encontradas em uma rede MLP:

- além da camada de entrada e da camada de saída, a MLP possui uma ou mais camadas ocultas que possibilitam à rede aprender tarefas complexas;
- cada neurônio da rede possui função de ativação não linear (não linearidade suave, ou seja, as funções são diferenciáveis em qualquer ponto), por exemplo, a função logística ou tangente hiperbólica;

- a MLP exibe um alto grau de conectividade, que é determinado pelas sinapses da rede.

O processo de treinamento das redes MLP utiliza o algoritmo *back-propagation* e é realizado por meio de duas fases distintas:

- propagação adiante (*forward*): os sinais $\{x_i\}$ de uma amostra do conjunto de treinamento são apresentados à rede e são propagados através das camadas até a obtenção das saídas da rede $\{y_i\}$. Nessa fase, são levados em consideração os pesos e limiares atuais, e estes não são alterados ao longo desta fase. Em seguida, os valores de saída são comparados aos respectivos valores esperados, os desvios (erros) relativos a cada uma das saídas e valores esperados são calculados, pois serão usados para o ajuste dos pesos e limiares;
- propagação reversa (*backward*): nesta fase, os valores dos erros encontrados na primeira fase são usados para o ajuste dos pesos sinápticos e limiares.

Assim, a cada iteração, novos ajustes dos pesos e limiares são realizados.

Pode-se dizer que cada neurônio oculto ou de saída de uma rede MLP é projetado para realizar dois cálculos (15):

- o cálculo do sinal funcional de saída de um neurônio, expresso como uma função não linear do sinal de entrada e dos pesos sinápticos associados a este neurônio;
- o cálculo de uma estimativa do vetor gradiente, que será usado na fase de propagação reversa.

Descrição do algoritmo backpropagation

A seguir, serão apresentadas as equações envolvidas nos cálculos do algoritmo *backpropagation* baseadas no livro de Silva, Spatti e Flauzino (64).

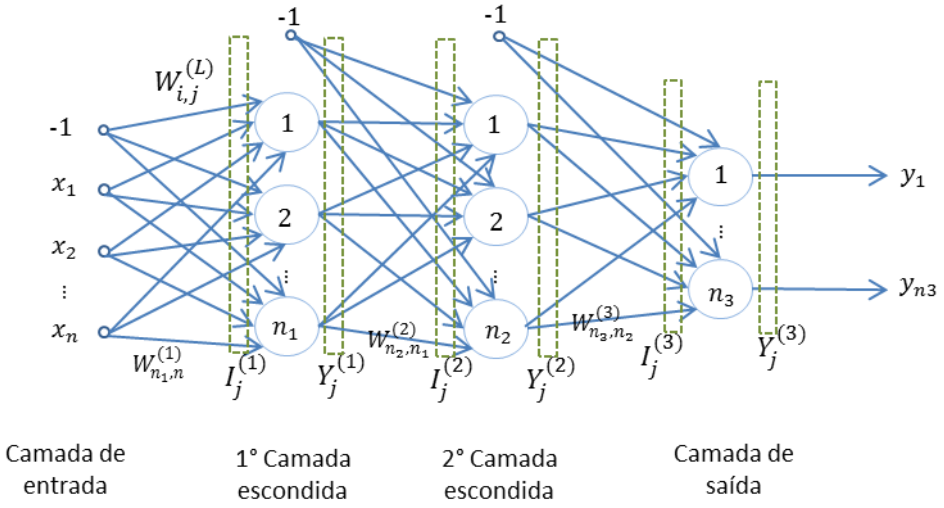

Figura 18 – Notação para algoritmo *backpropagation.* Fonte: SILVA, SPATTI e FLAUZINO (64).

A Figura 18 apresenta um esquemático da notação usada. Para cada neurônio $\{j\}$ em uma dada camada $\{L\}$ e cada neurônio $\{i\}$ da camada $\{L-1\}$, tem-se:

- $W_{j,i}^{(L)}$ são as matrizes de pesos sinápticos;
- $I_{j,i}^{(L)}$ são os vetores que representam as entradas ponderadas em relação ao *j*-ésimo neurônio da camada L e são definidos por:

$$I_j^{(L)} = \sum_{i=0}^{n} W_{j,i}^{(L)} \cdot x_i = W_{j,0}^{(L)} \cdot x_0 + W_{j,1}^{(L)} \cdot x_1 + \ldots + W_{j,n}^{(L)} \cdot x_n \ ; \tag{68}$$

- $Y_{j,i}^{(L)}$ são os vetores que representam a saída do *j*-ésimo neurônio da camada L e são definidos por:

$$Y_j^{(L)} = \varphi\left(I_j^L\right). \tag{69}$$

O desvio (erro quadrático) entre a resposta produzida pelos neurônios da camada de saída da rede (no caso a camada de número 3 contendo n_3 neurônios) e os valores desejados $d_j(k)$, para a *k*-ésima amostra, neste caso, é dado por:

$$E(k) = \frac{1}{2}\sum_{j=1}^{n_3}\left(d_j(k) - Y_j^{(3)}(k)\right)^2. \tag{70}$$

Para um conjunto de amostras de treinamento contendo p amostras, o erro quadrático médio é dado por:

$$E_M = \frac{1}{p}\sum_{k=1}^{p} E(k). \tag{71}$$

A descrição do processo de ajuste dos pesos sinápticos para o *backpropagation* é apresentada no apêndice B.

No caso da aplicação de redes MLP para classificação de padrões não linearmente separáveis, a rede deve ser também não linear. Cada neurônio produzirá um hiperplano no espaço de decisões, e a combinação desses hiperplanos é ajustada iterativamente no processo de aprendizagem de modo a separar os padrões.

A topologia da rede MLP a ser projetada depende da aplicação específica em que esta será empregada.

2.6.3 *Toolbox* de redes neurais artificiais do Matlab

O software Matlab possui o *Neural Network Toolbox*, que disponibiliza funções para modelagem de redes neurais e suporta diferentes tipos de arquiteturas de rede. Utilizando-se este *toolbox,* é possível criar uma rede neural e treiná-la (65-66).

Este *toolbox* possui ferramentas que também possibilitam visualizar e simular as redes; além disso, algumas ferramentas para avaliação do desempenho da rede também estão disponíveis, dentre elas a exibição de gráficos de regressão, que medem a correlação entre as saídas da rede e os valores esperados e histogramas de erros.

Uma forma gráfica desse *toolbox* também está disponível, sendo uma alternativa que facilita o uso dessas ferramentas para iniciantes.

3 TELESCÓPIOS REFLEXIVOS COM DOIS ELEMENTOS – ANÁLISE DE SENSIBILIDADE AO ALINHAMENTO

3.1 Introdução

Neste capítulo, serão apresentadas uma breve revisão sobre os tipos mais comuns de telescópios reflexivos com dois espelhos e uma configuração específica que revela uma característica peculiar: baixa sensibilidade a descentralização do espelho secundário (5).

Serão apresentadas algumas análises comparativas de sensibilidade ao alinhamento entre as diferentes configurações. Os erros de alinhamento avaliados foram inclinação e descentralização relativos entre os espelhos.

Um estudo detalhado do comportamento da configuração que apresenta baixa sensibilidade também será apresentado.

A abordagem usada na descrição e estudo dessa configuração especial baseou-se no trabalho prévio de Sasián (3) para sistemas plano-simétricos, apresentado no capítulo 2.

O trabalho descrito a seguir encontra-se publicado em: Optics Express, Vol. 21, Issue 6, pp. 6851-6865 (2013) (67).

3.2 Telescópios reflexivos com dois espelhos – descrição de algumas configurações

Telescópios com dois espelhos possuem um papel importante em Astronomia. No caso das configurações *Gregorian* e *Cassegrain*, o espelho primário é um paraboloide que produz uma imagem sem aberração em seu ponto focal sobre o

eixo óptico. O espelho secundário é posicionado de forma que seu foco coincida com o do espelho primário. Se o espelho secundário está localizado atrás do foco do primário, ou seja, há um ponto de foco entre os dois espelhos, o sistema é chamado de *Gregorian telescope* e possui distância focal negativa (Figura 19,a). Se o espelho secundário está localizado antes do foco do espelho primário, o telescópio é do tipo *Cassegrain* com distância focal positiva (Figura 19,b) (17).

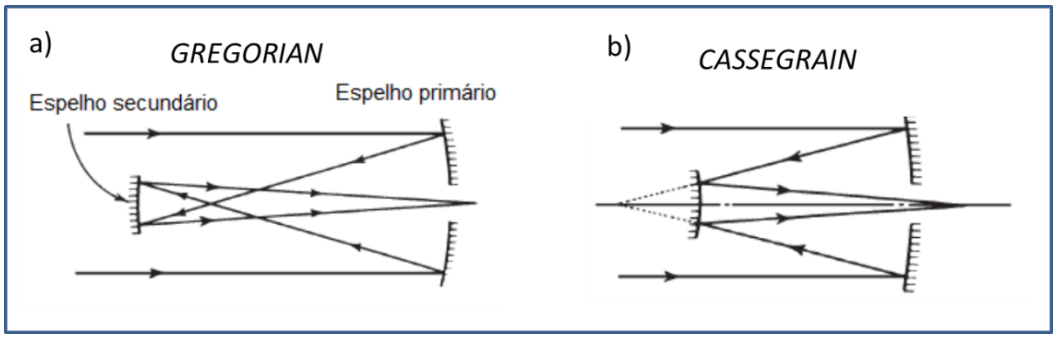

Figura 19 – Telescópios com dois espelhos: a) Gregorian e b) Cassegrain. Fonte: Adaptada de SMITH (17).

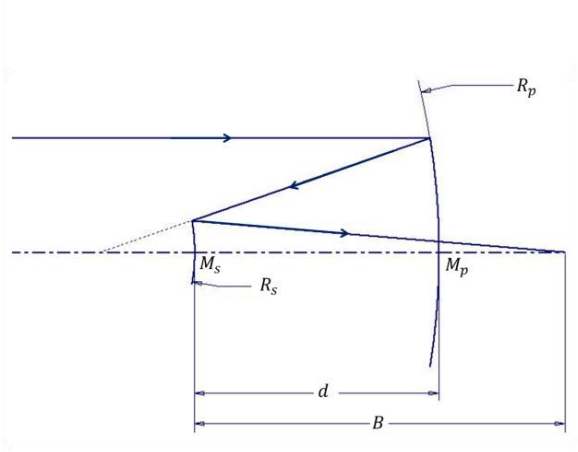

Figura 20 – Esquema ilustrativo das nomenclaturas adotadas Fonte: SCADUTO et al. (67).

Quando conhecidas a distância focal (f), a distância de foco posterior (B) e a separação entre os espelhos (d), pode-se encontrar os valores dos raios de curvaturas de cada um dos espelhos:

$$R_p = \frac{2df}{B-f}, \tag{72}$$

$$R_s = \frac{2dB}{B+d-f}\ , \tag{73}$$

Onde R_p é o raio de curvatura do espelho primário e R_s é o raio de curvatura do espelho secundário (Figura 20).

A constante cônica do espelho primário (k_p) pode ser ajustada de acordo com a constante cônica do secundário (k_s):

$$k_p = \left(\frac{B}{f}\right)\cdot\frac{(B+d-f)}{(B-f)^3}\cdot\left[(B+d-f)^2 k_s + (f+d-B)^2\right] - 1 \ . \tag{74}$$

Quando o espelho primário é uma parábola (constante cônica igual a -1), o telescópio é conhecido como *classical Gregorian* ou *classical Cassegrain*. Contudo, o espelho primário pode ter qualquer forma. A forma do secundário deve ser calculada de maneira que satisfaça condições específicas relacionadas às aberrações de terceira ordem do sistema. As configurações *classical Gregorian* e *classical Cassegrain* são obtidas se ambos os espelhos são independentemente corrigidos para aberração esférica de terceira ordem, o que leva a:

$$k_p = R_p^3\frac{(f-B)^3}{8d^3f^3} = -1, \tag{75}$$

$$k_s = R_s^3\frac{(f-d-B)(f+d-B)^2}{8d^3B^3}. \tag{76}$$

A configuração conhecida como *Ritchey-Chrétien* (RC) é corrigida tanto para aberração esférica como para coma de terceira ordem em todo o campo e possui constante cônica do espelho secundário dada por (17):

$$k_s = R_s^3\frac{\left[2f(B-f)^2 + (f-d-B)(f+d-B)(d-f-B)\right]}{8d^3B^3}. \tag{77}$$

Enquanto o campo de visada de telescópios *Cassegrain* é limitado por coma, que nesta configuração é nulo apenas sobre o eixo óptico, o campo de visada de telescópios RC é limitado por astigmatismo, com dependência quadrática do campo, uma vez que aberração esférica e coma são corrigidos em todo o campo nesta configuração (20, 68).

Outra configuração, conhecida como *Dall-Kirkham,* possui a vantagem de ter um espelho secundário esférico, de fácil fabricação. O telescópio *Dall-Kirkham* invertido possui o espelho primário esférico.

As equações fornecidas nesta seção funcionam para correção das aberrações esférica e de coma (para o RC) de terceira ordem. Ordens mais altas podem ser corrigidas, por exemplo, pelo ajuste dos termos asféricos de quarta a oitava ordem.

3.3 Alinhamento de sistemas reflexivos com dois elementos

Em sistemas reflexivos com dois elementos, o espelho primário pode ser considerado como a referência do sistema, uma vez que seus erros de posicionamento levarão a erros de apontamento do telescópio (6), mas não introduzirão nenhum tipo de aberração no sistema, se o espelho secundário estiver alinhado em relação a ele. Por esse motivo, os erros de alinhamento considerados aqui ocorrem no espelho secundário, em relação ao espelho primário.

O espelho secundário pode estar descentralizado ou inclinado. Erros de posicionamento axial não serão considerados nas análises apresentadas ao longo deste capítulo.

Um elemento estará inclinado, se o seu eixo não for paralelo ao eixo óptico do sistema e estará descentralizado, se seu vértice não estiver sobre o eixo óptico do sistema.

Conforme apresentado no capítulo 2, esses erros de alinhamento levam ao aparecimento de novos termos de aberrações com dependências com o campo diferentes das dependências dos termos que podem existir em sistemas com simetria axial.

Em sistemas com dois espelhos, erros de alinhamento do espelho secundário causam quebra de simetria axial e podem gerar principalmente astigmatismo linear e coma uniforme, sendo o último o mais importante neste tipo de sistema (5). Essas aberrações aparecem em sistemas plano-simétricos. Quando apenas inclinação ou apenas descentralização está presente no espelho secundário de sistemas com dois espelhos, o sistema passa a ser plano-simétrico e, por isso, a teoria de Sasián (3) apresentada no capítulo 2 foi usada nos desenvolvimentos apresentados neste capítulo.

Os efeitos causados por erros de alinhamento nos campos de aberrações de terceira ordem de sistemas com dois espelhos foram apresentados no trabalho de Schmid (20) com detalhes. O autor baseou seus estudos na Teoria de Aberração de Onda Nodal (*Nodal Aberration Theory*) proposta por Thompson (2) e apresentada no capítulo 2 desta tese.

Por ser bastante utilizado, diversas técnicas de alinhamento foram propostas para este tipo de sistema e baseiam-se na correção, no eixo óptico, das aberrações coma e astigmatismo (5-6, 20, 23, 41, 69).

Vários métodos de alinhamento são baseados na eliminação da aberração coma no eixo óptico. Para alguns tipos de telescópios *Cassegrain*, com *f/#* lento[j] e pequenos campos de visada, esses métodos podem ser suficientes. Contudo, para telescópios aplanáticos (sistemas corrigidos para aberração esférica e coma), como os RC, com amplo campo de visada, alinhamentos baseados na correção do coma e astigmatismo no eixo óptico podem não ser suficientes, visto que o astigmatismo pode apresentar dois pontos em que seu valor é nulo, possibilitando que o sistema ainda permaneça desalinhado mesmo estando corrigido no eixo óptico (20, 22, 41).

Correções do astigmatismo podem ser realizadas por meio de medidas do valor desta aberração em pontos fora do eixo (69).

3.4 Sensibilidade ao alinhamento

Nesta seção, será apresentada uma análise da sensibilidade a erros de alinhamento de diferentes configurações de sistemas reflexivos com dois espelhos do tipo *Cassegrain*.

As configurações avaliadas possuem EFL de 3750 mm, BFL de 1200 mm, separação entre os espelhos de 800 mm e diferentes constantes cônicas. As equações (72) e (73) possibilitam o cálculo dos raios de curvatura dos espelhos primário (R_p=-2352,94 mm) e secundário (R_s=-1097,14 mm). A constante cônica de

[j] Os termos “rápido” e “lento” estão relacionados ao *f*-number (*f/#,* dado pela razão da distância focal efetiva pela abertura) de um sistema óptico. Sistemas com uma abertura grande (e *f/#* pequeno) são chamados de rápidos e sistemas com pequena abertura (e *f/#* grande) são chamados de lentos (17).

um dos espelhos foi variada em um determinado intervalo e, para cada valor, foi encontrada a constante cônica ideal para o outro espelho, de forma que a condição de aberração esférica nula fosse obedecida, ou seja, utilizou-se a equação (74) (Tabela 6).

Tabela 6 – Combinação de constantes cônicas dos espelhos.

k_p	k_s
-1,00	-3,66
-0,90	-2,70
-0,80	-1,73
-0,70	-0,76
-0,60	0,20
-0,50	1,17
-0,48	1,36
-0,44	1,75
-0,42	1,94
-0,40	2,14
-0,38	2,33
-0,36	2,52
-0,34	2,72
-0,30	3,10
-0,20	4,07
-0,10	5,04
0,00	6,00

Fonte: SCADUTO et al. (67).

O efeito de perturbações aplicadas no espelho secundário foi avaliado por meio do valor do erro RMS (*root mean square*) no eixo. A Figura 21 exibe a sensibilidade (dada pelo valor RMS) em função dos desvios de alinhamento do espelho secundário para perturbações do tipo descentralização e inclinação respectivamente. Diferentes configurações apresentam diferentes sensibilidade a descentralizações no espelho secundário. Porém, observa-se que as configurações analisadas apresentam sensibilidade a inclinações do espelho secundário similares.

Esses resultados indicam que há uma configuração específica que é menos afetada por erros de alinhamento do tipo descentralização e possui sensibilidade similar às outras no que diz respeito a erros de alinhamento do tipo inclinação. No caso do sistema analisado, a configuração cujo espelho primário possui constante cônica de -0,42 parece ser a de menor sensibilidade.

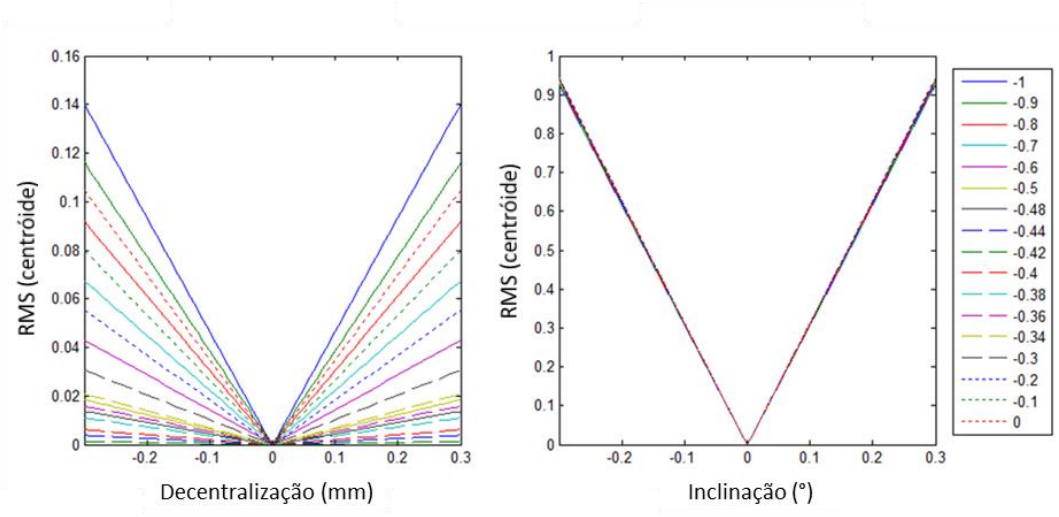

Figura 21 – O gráfico à esquerda apresenta a sensibilidade a perturbações do tipo descentralização para diferentes configurações com mesmos f, d e B. O gráfico à direita exibe a sensibilidade a perturbações do tipo inclinação para as mesmas configurações. A legenda apresenta o valor da constante cônica do espelho primário de cada configuração analisada. Fonte: Adaptado de SCADUTO et al. (67).

Por possuir essa característica peculiar, uma nova comparação foi realizada entre sistemas que apresentam $f = 3750$ mm, $B = 1200$ mm e $d = 800$ mm. Três configurações foram escolhidas: uma do tipo *classic Cassegrain* (com constante cônica do primário igual a -1 e ambos os espelhos corrigidos para aberração esférica), uma do tipo *Ritchey-Chrétien* (corrigido para aberração esférica e coma-linear) e a configuração menos sensível (*Less-sensitive* - LS) (5) específica para os parâmetros apresentados anteriormente. As características de cada uma delas são mostradas na Tabela 7. Esta tabela também apresenta os valores do coma linear residual (W_{13100}) e do astigmatismo quadrático (W_{22200}) para um semicampo de visada de 0,25°.

Em todas as configurações analisadas, o *STOP* está localizado no espelho primário, que é a referência do sistema; sendo assim, a sensibilidade das diferentes configurações foram avaliadas em função da presença de desalinhamentos no espelho secundário. Assim como na análise anterior, perturbações do tipo descentralização e inclinação foram avaliadas.

Tabela 7 – Constantes cônicas, termos asféricos do espelho primário, coma linear residual e astigmatismo quadrático para as configurações analisadas.

Configurações	k_p	k_s	M_p **even asph. 4th**	M_p **even asph. 6th**	M_p **even asph. 8th**	W_{13100} **(waves) (λ=0,55μm)**	W_{22200} **(waves) (λ=0,55μm)**
Classic Cassegrain	-1,0	-3,7	3,52E-14	-8,43E-23	-7,53E-27	-2,1	0,7
Ritchey-Chrétien (RC)	-1,1	-4,6	-3,11E-14	-1,02E-20	-2,08E-27	0,098	0,8
Less-Sensitive (LS)	-0,4	1,9	2,35E-13	3,13E-19	2,09E-25	-15,9	0,1

Fonte: Adaptado de SCADUTO et al. (67).

Para que fosse possível comparar as configurações, os polinômios de Zernike foram analisados em cada situação e os perfis das aberrações coma e astigmatismo, que são as mais afetadas por erros de alinhamento nestes sistemas, foram obtidos por meio das equações[k]:

$$Coma_{total} = \sqrt{Z_7^2 + Z_8^2}, \tag{78}$$

$$Astigmatismo_{total} = \sqrt{Z_5^2 + Z_6^2}, \tag{79}$$

A Figura 22 apresenta os gráficos de coma total (conforme calculado na equação (78)) em todo o campo para as três configurações analisadas, da esquerda para a direita: *classical Cassegrain*, RC e LS, no estado alinhado, com descentralização e com inclinação no espelho secundário.

Nas duas primeiras configurações, as curvas de coma total de terceira ordem são claramente afetadas tanto para descentralização como para inclinação. No caso da configuração menos sensível, apenas inclinação altera o perfil da curva e, portanto, descentralizações no espelho secundário não geram termos adicionais para a aberração coma.

Para uma comparação entre os três sistemas, os coeficientes dos polinômios de Zernike (*Fringe*) (Z_5, Z_6, Z_7 e Z_8) foram extraídos para o estado alinhado e estados perturbados dos sistemas. Os coeficientes Z_5 e Z_6 estão associados à aberração astigmatismo e os coeficientes Z_7 e Z_8 estão associados à aberração coma de terceira ordem.

[k] De acordo com a notação de Wyant e Creath (44).

Os coeficientes Z_5 e Z_8 são os mais afetados quando perturbações do tipo inclinação (aplicadas no eixo *x*) estão presentes no sistema (Figura 24). Observa-se que as três configurações apresentam comportamentos similares.

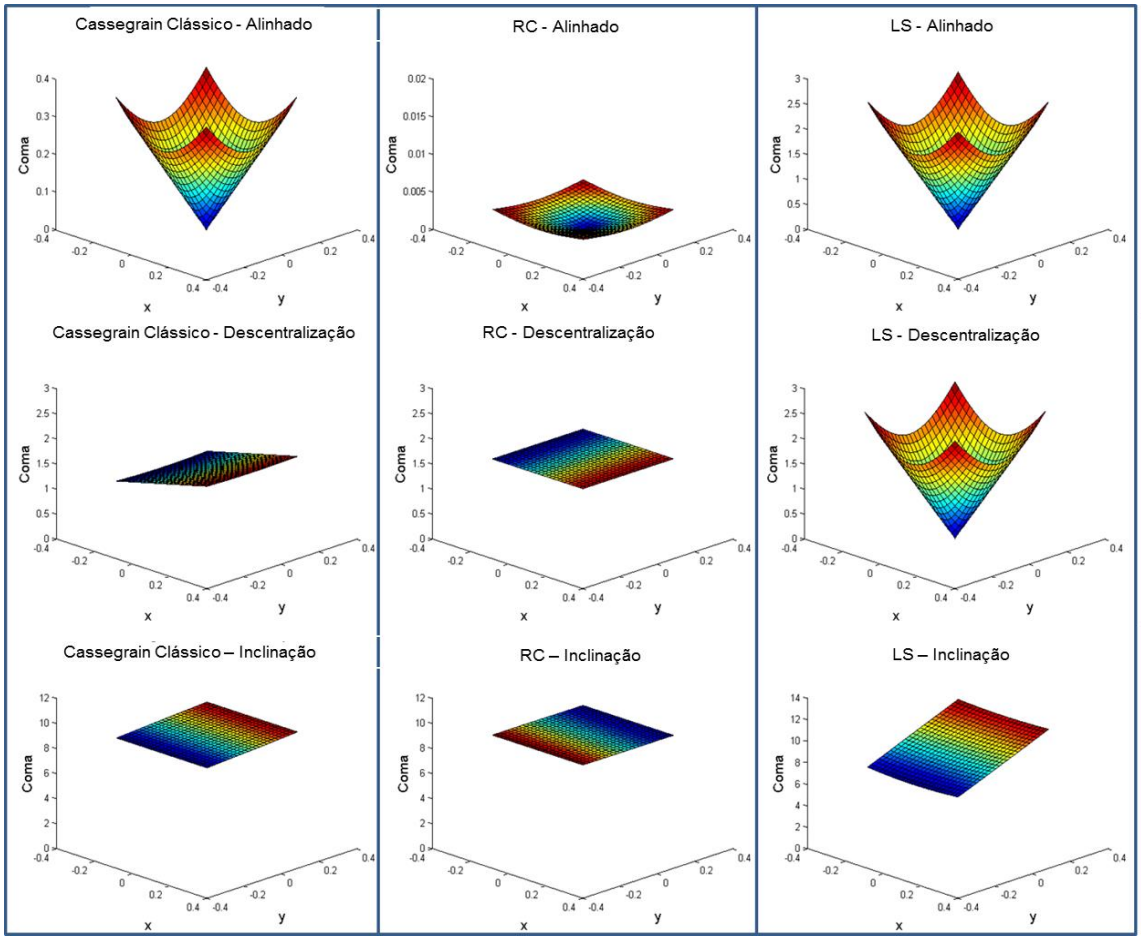

Figura 22 – Gráficos de coma total (a menos de um fator de escala) em todo o campo (waves, λ=0,55 μm) dado pela equação (78). A figura à esquerda apresenta curvas do coma em todo o campo para a configuração *classical Cassegrain*, no estado alinhado, com descentralização [mm] e com inclinação [°] no espelho secundário, respectivamente, de cima para baixo. Os mesmos gráficos são apresentados na coluna central para a configuração RC e na coluna à direita para a configuração com menor sensibilidade (LS).

O coeficiente Z_7 é o mais afetado quando perturbações do tipo descentralização são inseridas no eixo *x* do espelho secundário. A Figura 25 mostra que a configuração menos sensível não é afetada por coma originário de descentralização. Astigmatismo, contudo, ainda é afetado neste caso.

Conclui-se, assim, que para essa configuração, coma constante em todo o campo é nulo mesmo na presença de descentralizações no sistema.

O astigmatismo total, como pode ser observado na Figura 23, muda na presença tanto de descentralização como de inclinação no secundário em todas as configurações avaliadas.

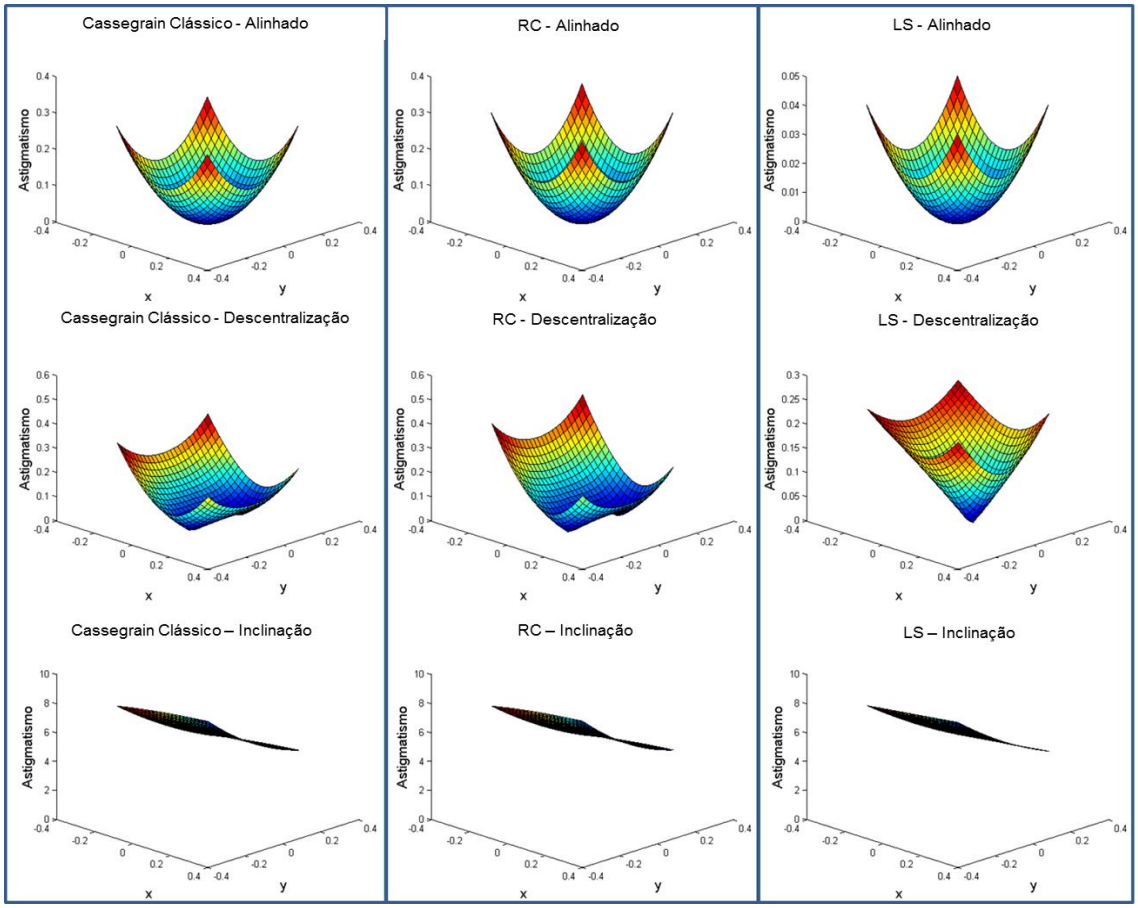

Figura 23 – Gráficos de astigmatismo total (a menos de um fator de escala) em todo o campo (waves, λ=0,55 μm) dado pela equação (79). A figura à esquerda apresenta curvas de astigmatismo em todo o campo para a configuração *classical Cassegrain*, no estado alinhado, com descentralização [mm] e com inclinação [°] no espelho secundário, respectivamente, de cima para baixo. Os mesmos gráficos são apresentados na coluna central para a configuração RC e, na coluna à direita, para a configuração com menor sensibilidade (LS).

Portanto, o que faz que a configuração LS seja menos sensível a erros de alinhamento do tipo descentralização é o fato de não haver inserção de coma constante no sistema quando esse tipo de desalinhamento está presente no sistema. Desta forma, para essa configuração, astigmatismo será a aberração dominante devido a desvios de alinhamento quando não houver desalinhamentos do tipo inclinação no sistema.

Esta configuração, contudo, ainda é bastante sensível à presença de inclinações no espelho secundário, que alteram tanto coma como astigmatismo em todo o campo, inclusive no eixo, fazendo com que os valores dessas aberrações no eixo deixem de ser zero em estados desalinhados.

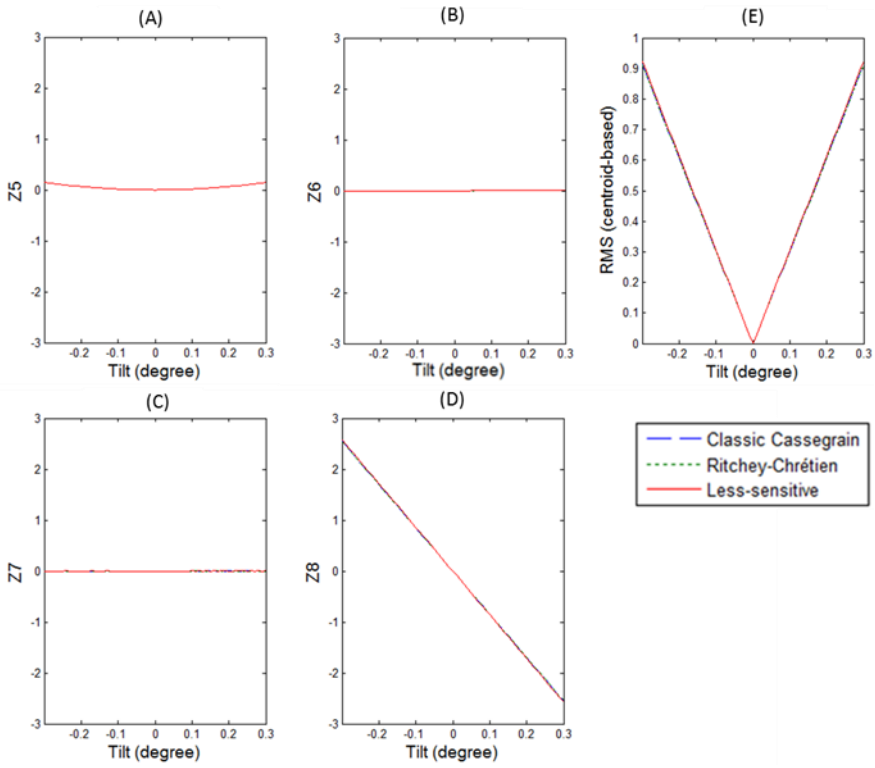

Figura 24 - Gráficos dos coeficientes dos polinômios de Zernike: A) Z_5 ; B) Z_6 ; C) Z_7 ; D) Z_8 e E) RMS (waves, λ=0.55 μm) em função da inclinação do espelho secundário para as configurações *classical Cassegrain*, RC e LS. Fonte: SCADUTO et al. (67).

3.5 A configuração de menor sensibilidade - LS

No capítulo 2, foram apresentadas as principais teorias de aberrações de onda em sistemas ópticos axialmente simétricos e sistemas desalinhados ou sem simetria axial.

As aberrações de Seidel não se aplicam quando há erros de alinhamento no sistema. As teorias apresentadas, desenvolvidas por Thompson (2) e Sasián (3), buscam explicar os efeitos que esses erros de alinhamento causam no sistema e fornecem equações para quantificá-los.

A teoria de Thompson pode ser aplicada a qualquer tipo de desalinhamento em qualquer direção.

A teoria de Sasián foi desenvolvida para sistemas com erros de alinhamento restritos aos que tornam sistemas axialmente simétricos em sistemas plano-simétricos e as equações fornecidas são aplicáveis a desalinhamentos do tipo inclinação. Essa abordagem tem a vantagem de fornecer uma forma simples e direta de identificar e entender o que os erros de alinhamento estão causando ao sistema. Essa teoria é aplicável a sistemas com elementos descentralizados desde que algumas adaptações sejam aplicadas às fórmulas.

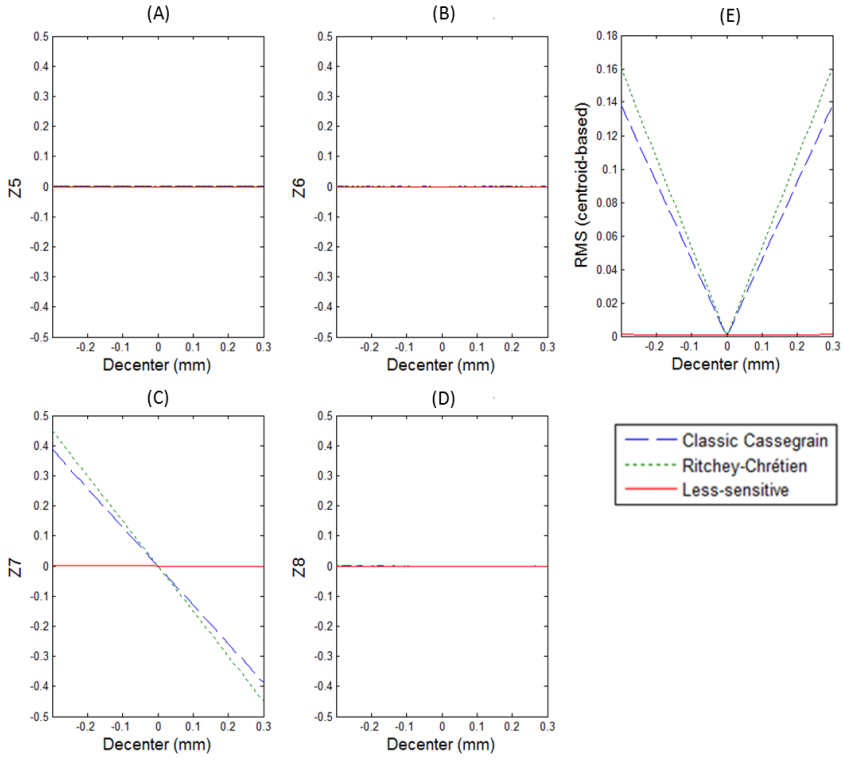

Figura 25 – Gráficos dos coeficientes dos polinômios de Zernike: A) Z_5 ; B) Z_6 ; C) Z_7; D) Z_8 e E) RMS (waves, λ=0.55 μm) em função de descentralização do espelho secundário para as configurações *classical Cassegrain*, RC e LS. Fonte: SCADUTO et al. (67).

A teoria de aberração de onda aplicável a sistemas plano-simétricos foi usada neste trabalho para explicar os efeitos causados pelos erros de alinhamento nos sistemas reflexivos analisados na seção anterior, em particular o sistema que apresenta menor sensibilidade a descentralização. As fórmulas propostas originalmente também foram modificadas para abranger elementos descentralizados.

De acordo com as teorias apresentadas, erros de alinhamento geram um termo adicional à aberração coma. Em sistemas alinhados simétricos axialmente, o termo do coma que pode estar presente (caso não seja corrigido em todo o campo, sendo assim nulo) é o coma com dependência linear com o campo. Quando erros de alinhamento ocorrem, um termo constante com o campo surge e deve ser adicionado ao primeiro.

A conclusão obtida na análise apresentada anteriormente foi que no sistema com menor sensibilidade a desalinhamentos, coma uniforme (W_{03001}) não é gerado quando descentralizações estão presentes no sistema, uma vez que o campo desta aberração não muda para este tipo de perturbação. Esse sistema, já discutido por Wilson (5), é abordado aqui de forma diferente: a teoria de Sasián para sistemas plano-simétricos é aplicada e, a partir dela, busca-se avaliar esse sistema peculiar e entender seu comportamento.

As contribuições de cada superfície para esse termo de aberração, de acordo com a teoria adotada, são apresentadas a seguir.

Quando um sistema possui superfícies asféricas, o *sag*[l] da superfície é dado pela soma das contribuições esférica e asférica:

$$z = z_{esferico} + z_{asferico}\,. \tag{80}$$

3.5.1 Contribuição esférica

Conforme apresentado no capítulo 2 e no apêndice A, em um sistema plano-simétrico a contribuição esférica para o coma uniforme de uma superfície inclinada por um ângulo l é dada pela equação:

$$W^{0}_{03001} = \sum_{i=1}^{i=j} \{J_{II}\}_i\,, \tag{81}$$

onde,

[l] Abreviação de *sagitta*. Descreve a altura de uma curva medida a partir da corda. Fonte: The Photonics Dictionary, Pittsfield: Laurin Publishing, p. D122, 2004.

$$J_{II} = -\frac{1}{2} n \operatorname{sen}(I) A \Delta\left(\frac{u}{n}\right) x. \tag{82}$$

Na equação (82), n é o índice de refração do meio. Para um sistema reflexivo com dois elementos no ar, $n_p = n'_s = 1$ e $n'_p = n_s = -1$. A grandeza A, conhecida como Invariante de Snell (5), é dada pela equação $A = ni = n\left(u + \frac{x\cos(I)}{R}\right)$ e Δ, chamada de *Abbe difference operator,* é dada pela equação $\Delta\left(\frac{u}{n}\right) = \frac{u'}{n'} - \frac{u}{n}$.

A contribuição esférica total no sistema com dois espelhos é dada pela soma das contribuições dos dois elementos:

$$W^o_{03001} = \left\{J^p_{II} + J^s_{II}\right\} = -\frac{1}{2}\left[n_p \operatorname{sen}(I_p) A_p \Delta\left(\frac{u}{n}\right)_p x_p + n_s \operatorname{sen}(I_s) A_s \Delta\left(\frac{u}{n}\right)_s x_s\right], \tag{83}$$

onde u é a inclinação do raio marginal e x é a altura deste raio nesta superfície.

As equações de (84) a (87) apresentam os valores das grandezas relacionadas a cada uma das superfícies deste sistema.

$$A_p = \frac{m}{2f} x_p \cos(I_p), \tag{84}$$

$$A_s = \frac{m}{f}\left[1 - \frac{(m+1)}{2m}\cos(I_s)\right] x_p, \tag{85}$$

$$\Delta\left(\frac{u}{n}\right)_p = \frac{m}{f} x_p, \tag{86}$$

$$\Delta\left(\frac{u}{n}\right)_s = -\frac{m}{f}\left(1 + \frac{1}{m}\right) x_p. \tag{87}$$

Da equação (83), tem-se:

$$W^o_{03001} = -\frac{1}{2}\left[\frac{1}{2}\left(\frac{m}{f}\right)^2 \operatorname{sen}(I_p)\cos(I_p) + \frac{Bm}{f^3}\operatorname{sen}(I_s)(m+1)\left[1 - \frac{(m+1)}{2m}\cos(I_s)\right]\right] x_p^3 , \tag{88}$$

onde $m = f / f'_1$, e f'_1 é a distância focal efetiva do espelho primário.

Como o espelho primário é a referência do sistema, este não sofre nenhum tipo de erro de alinhamento e, portanto, $I_p = 0$, e, consequentemente, $\operatorname{sen}(I_p) = 0$. Dessa

forma, apenas o espelho secundário contribui para a porção esférica do coma uniforme na presença de erros de alinhamento do tipo inclinação:

$$W_{03001}^{o} = -\frac{Bm}{2f^3}\,\mathrm{sen}(I_s)(m+1)\left[1-\frac{(m+1)}{2m}\cos(I_s)\right]x_p^3\ . \tag{89}$$

A equação anterior pode ser adaptada para que os efeitos de descentralização no espelho secundário sejam avaliados.

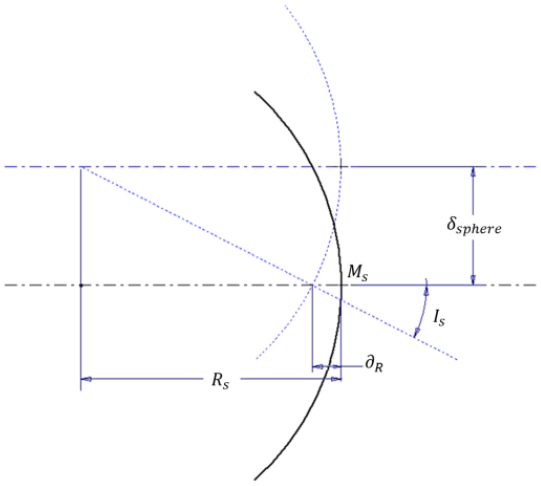

Figura 26 – Descentralização em uma superfície esférica. Fonte: SCADUTO et al. (67).

Quando uma superfície esférica está descentralizada, o seu vértice está deslocado em relação ao eixo óptico do sistema (Figura 26). O efeito desse deslocamento, no ponto onde a superfície cruza o eixo óptico, pode ser considerado como um deslocamento ($\delta_{esferico}$) do centro de curvatura da superfície. O mesmo efeito ocorre se a superfície estiver inclinada (I_s) em torno de seu vértice. Assumindo-se que o deslocamento axial da superfície (∂_R) é desprezível, o deslocamento $\delta_{esferico}$ pode ser descrito em termos do ângulo I_s como:

$$\mathrm{sen}(I_s) = \frac{\delta_{esferico}}{R_s}. \tag{90}$$

Sendo I_s pequeno, outra aproximação válida é $\cos(I_s) = 1$, e a contribuição esférica para o coma uniforme quando descentralizações estão presentes no espelho secundário é dada por:

$$W_{03001}^{o} = -\frac{1}{8}\left(\frac{x_p}{f}\right)^3 (m+1)^2 (m-1)\,\delta_{esferico}\ . \tag{91}$$

Assim, a equação (89) é válida para perturbações do tipo inclinação em sistemas com dois espelhos e a equação (91) vale para perturbações do tipo descentralização.

3.5.2 Contribuição asférica

No caso da contribuição asférica, quando há erros de alinhamento do tipo inclinação no sistema, Sasián considera que a contribuição asférica é dada por incrementos ΔW apresentados na Tabela 30 do apêndice A. Para o coma uniforme, esta contribuição é dada por:

$$W_{03001}^{*} = \Delta W_{03001} = \beta\Delta\left[n\cos(I)\right]x^3. \tag{92}$$

Um fato interessante no termo asférico do coma uniforme que surge por causa de erros de alinhamento do tipo inclinação é o fato de ele ser independente das constantes cônicas dos espelhos (5). No que diz respeito a este tipo de desalinhamento, a sensibilidade depende do raio de curvatura da superfície do elemento. Assim, independentemente da configuração adotada, sistemas com mesmo f, d e B terão a mesma contribuição asférica para esta aberração.

Esse fato explica por que as diferentes configurações analisadas apresentam sensibilidade diferente na presença de elementos descentralizados, mas a sensibilidade à inclinação praticamente não muda de uma configuração para outra.

No caso de desalinhamentos devido a descentralização do espelho secundário, a contribuição da porção asférica para o coma de terceira ordem pode ser obtida por meio da equação do *sag* de uma superfície asférica:

$$z = \frac{(x^2+y^2)}{2R} + (k+1)\frac{(x^2+y^2)^2}{8R^3}. \tag{93}$$

A contribuição asférica para o *sag* da superfície do espelho secundário é dada por:

$$z_{asferico} = k_s \frac{(x_s^2 + y_s^2)^2}{8R_s^3}. \tag{94}$$

A frente de onda é dada por:

$$W^* = k_s \frac{\left(x_s^2 + y_s^2\right)^2}{8R_s^3} \Delta(n)_s. \tag{95}$$

Se o feixe está descentralizado de $\delta_{asferico}$ no eixo *x*, a frente de onda torna-se:

$$W^*_{decentr.} = k_s \frac{\left((x_s + \delta_{asferico})^2 + y_s^2\right)^2}{8R_s^3} \Delta(n)_s. \tag{96}$$

E a contribuição asférica para o coma uniforme de terceira ordem em um sistema de dois espelhos é dada por:

$$W^*_{03001} = k_s \frac{4x^3 \delta_{asferico}}{8R_s^3} \Delta(n)_s. \tag{97}$$

Assim, tem-se:

$$W^*_{03001} = \frac{1}{8}\left(\frac{x_p}{f}\right)^3 (m+1)^3 k_s \delta_{asferico}. \tag{98}$$

3.5.3 Termo total da aberração coma uniforme para sistemas reflexivos com dois elementos

O termo total do coma uniforme, que é gerado devido a erros de alinhamento nos elementos ópticos do sistema, quando há descentralização do espelho secundário é dado pela soma das contribuições esférica e asférica:

$$W_{03001} = -\frac{1}{8}\left(\frac{x_p}{f}\right)^3 (m+1)^2 \left[(m-1)\delta_{esferico} - (m+1)k_s \delta_{asferico}\right]. \tag{99}$$

Para erros de alinhamento iguais das porções esférica e asférica, a equação fica reduzida à:

$$W_{03001} = -\frac{1}{8}\left(\frac{x_p}{f}\right)^3 (m+1)^2 \left[(m-1)-(m+1)k_s\right]\delta. \quad (100)$$

No sistema que apresenta menor sensibilidade, o coeficiente de coma uniforme deve ser nulo, como foi observado nos gráficos apresentados na Figura 22, para que o coma em todo o campo se mantenha inalterado na presença de descentralização no espelho secundário. Assim, igualando a zero a equação anterior, tem-se:

$$\left[(m-1)-(m+1)k_s\right] = 0, \quad (101)$$

o que leva a:

$$k_s = \frac{(m-1)}{(m+1)}. \quad (102)$$

Como $m = \frac{f}{f_1'} = \frac{(B-f)}{d}$, a equação para k_s pode ser reescrita como:

$$k_s = \frac{B-f-d}{B-f+d}. \quad (103)$$

A equação (103) corresponde a uma solução para o valor da constante cônica do espelho secundário para que o sistema tenha baixa sensibilidade a descentralização, já que para esse valor de constante cônica o termo de coma uniforme é nulo e o valor dessa aberração em todo o campo se mantém inalterado. Contudo, o sistema ainda apresenta alterações no campo do astigmatismo, e a sensibilidade a inclinações no espelho secundário não muda em relação a qualquer outra combinação de constantes cônicas.

Como o valor da constante cônica do espelho primário depende da constante do espelho secundário, dados f, d e B, o sistema LS possui solução completa. Os valores encontrados por meio da equação (103) coincidem, até a primeira casa decimal, com os valores obtidos por meio das análises de sensibilidade realizadas anteriormente (Figura 27), onde sistemas semelhantes, com diferentes constantes cônicas, foram perturbados e a sua sensibilidade ao alinhamento foi avaliada.

A sensibilidade do sistema a descentralizações do espelho secundário é menor à medida que o valor da constante cônica do espelho secundário se aproxima do valor obtido pela equação (103).

3.6 O ponto livre de coma

Sistemas reflexivos com dois componentes possuem dois pontos neutros: o ponto neutro para apontamento (*neutral point for pointing*), localizado sempre no centro de curvatura do espelho secundário, independente de sua forma; e o ponto livre de coma (*coma-free point),* localizado em algum ponto sobre o eixo óptico (5).

Coma gerado por rotação e por translação do espelho secundário se cancelam e, portanto, se o espelho secundário é rotacionado em torno do ponto livre de coma, nenhum coma é adicionado ao sistema.

Esses pontos são usados para o alinhamento desses sistemas, em especial o ponto livre de coma, já que torna possível a correção do astigmatismo sem inserir coma (5-6, 69).

Igualando a zero a equação dada por Wilson (5) para o coma total, pode-se encontrar uma equação para o ponto livre de coma Z_{CFP} em um sistema reflexivo com dois elementos:

$$Z_{CFP} = \frac{2f(m-1)R_A}{(m+1)^2\left[\frac{(m-1)}{(m+1)} - k_s\right]}. \tag{104}$$

Na equação (104),

$$R_A = \left(1 + {}^{2d}\!/_{R_p}\right). \tag{105}$$

A mesma solução foi obtida por Ren et al. (42), que derivaram esta equação partindo da teoria de aberração de onda nodal. Esta solução também pode ser deduzida partindo da equação apresentada anteriormente como solução do termo total de coma uniforme em um sistema com descentralização.

Quando k_s do sistema LS é substituido na equação para o ponto livre de coma, o resultado é $Z_{CFP} = \infty$, ou seja, nestes sistemas o ponto livre de coma está localizado no infinito, o que explica o fato de descentralizações do espelho secundário não alterar o campo do coma.

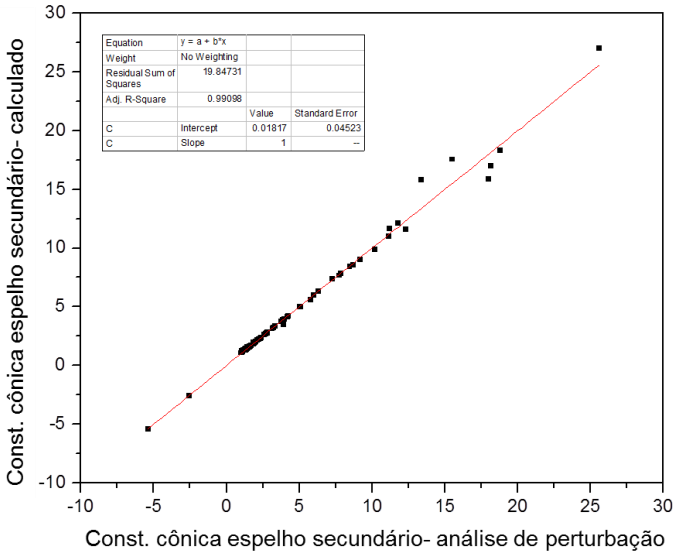

Figura 27 – Gráfico dos valores de k_S calculados pela equação (103) versus valores encontrados nas análises de perturbação para sistemas com diferentes combinações de *f*, *B* e *d*.

A Figura 28 apresenta o comportamento do posicionamento do ponto livre de coma em função do valor da constante cônica do espelho secundário para dois sistemas de mesma distância focal f e valores distintos de d e B.

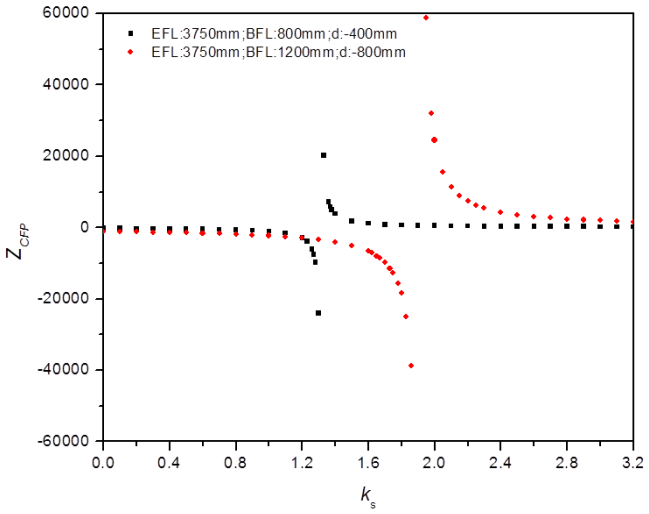

Figura 28 – Gráfico da posição do ponto livre de coma Z_{CFP} em função da constante cônica do espelho secundário em sistemas de dois espelhos. Fonte: SCADUTO et al. (67).

O gráfico mostra que Z_{CFP} apresenta uma singularidade em um valor específico de k_s, que corresponde ao valor para o qual a configuração do sistema apresentará menor sensibilidade a erros de alinhamento.

3.7 Conclusões deste capítulo

Os estudos apresentados neste capítulo mostram que a sensibilidade ao alinhamento de sistemas com dois espelhos depende da configuração adotada. No caso de sistemas com mesma distância focal, distância de foco posterior e separação entre os espelhos, que satisfaçam a condição de aberração esférica total do sistema nula, restringindo as curvaturas dos espelhos de modo que o sistema se enquadre em telescópios do tipo *Cassegrain* (com espelho primário côncavo e secundário convexo), os valores das constantes cônicas dos espelhos estão diretamente relacionados com a sua sensibilidade quanto a descentralização do espelho secundário. No caso de inclinações, a sensibilidade de diferentes configurações é praticamente igual, já que não depende da constante cônica dos espelhos.

Para uma configuração específica, chamada aqui de sistema com menor sensibilidade, ou LS, a constante cônica do espelho secundário faz que o ponto livre de coma esteja localizado muito longe do sistema, tendendo ao infinito. Essa solução faz com que a contribuição da parte esférica das superfícies ao coma uniforme (no caso aqui, da superfície do espelho secundário, já que o espelho primário é a referência e o *STOP* está localizado nele) seja completamente compensada pela contribuição da parte asférica da superfície, resultando em um termo nulo de coma uniforme. Dessa forma, para essa configuração, a constante cônica escolhida faz com que nenhum coma seja adicionado ao sistema e, portanto, o valor do coma ao longo de todo o campo não muda para estados descentralizados do sistema. Com isso, o termo de aberração mais importante gerado por erros de alinhamento em sistemas deste tipo deixa de existir, tornando o astigmatismo a aberração dominante devido a descentralização.

Apesar de não ser gerado coma uniforme quando há descentralização no espelho secundário, o sistema, em seu estado alinhado, apresenta coma linear.

O coma ao longo do campo não é alterado por descentralização e, portanto, o seu ponto de zero se mantém sobre o eixo óptico.

Como elementos descentralizados ocasionam astigmatismo linear no sistema, o valor do astigmatismo deixa de ser nulo no eixo, e o seu ponto de zero se move para outra posição no campo.

Esse fato pode facilitar o alinhamento desse tipo de sistema, já que a alteração do valor do coma no eixo é consequência de desalinhamentos do tipo inclinação no espelho secundário. Assim, alterando-se apenas a inclinação do espelho secundário pode-se corrigir o sistema até que o coma no eixo seja nulo.

Outro fato interessante é que essa configuração é compatível com uma estrutura para telescópios conhecida como Treliça de Serrurier (*Serrurier truss*). Ela mantém os eixos dos espelhos sempre paralelos, o que significa que não insere inclinações no sistema. Descentralizações que eventualmente podem ocorrer não causam graves degradações no sistema, uma vez que coma uniforme não é gerado.

4 METODOLOGIA PARA ALINHAMENTO DE SISTEMAS ÓPTICOS

4.1 Introdução

Esta seção apresenta a abordagem e as hipóteses adotadas neste trabalho relativos à proposta de metodologia para alinhamento óptico baseado na análise das aberrações de frente de onda do sistema óptico. Os detalhes do problema que se busca solucionar serão reapresentados visando justificar a abordagem proposta.

4.2 Descrição do problema

Na revisão bibliográfica, apresentada no Capítulo 2, foram descritas as aberrações que podem estar presentes em sistemas ópticos simétricos: aberração esférica, coma, astigmatismo, curvatura de campo e distorção.

No caso de sistemas com simetria rotacional, estas aberrações aparecem com uma dependência específica do campo, porém, quando há erros de alinhamento no sistema óptico, sejam eles devidos a erros de posicionamento dos elementos durante o processo de alinhamento ou mesmo desalinhamentos inseridos no sistema de forma intencional, não há a criação de novas aberrações, mas as aberrações já existentes ganham novas dependências com o campo. Nestes casos, o sistema deixa de ter simetria rotacional.

Desta forma, os termos das aberrações geradas por erros de alinhamento em um sistema óptico são distintos dos termos das aberrações presentes quando este se encontra no estado alinhado. Os novos termos gerados por desalinhamentos apresentam diferentes dependências com o campo, possuindo sempre dependências de ordens mais baixas que as já existentes (2).

Estas novas dependências com o campo alteram a frente de onda do sistema óptico e, por este motivo, esta frente de onda contém informações a respeito dos desalinhamentos presentes nos elementos do sistema óptico.

No Capítulo 3, foram discutidos, com mais detalhes, os efeitos causados por descentralização e inclinação nas aberrações de sistemas reflexivos com dois elementos. Estes efeitos ocorrem também em sistemas refrativos e da mesma forma que são alterados coma e astigmatismo, descentralização e inclinação alteram também as demais aberrações presentes no sistema, com exceção da aberração esférica de terceira ordem, que não possui dependência com o campo (2-3).

Em sistemas ópticos reais, a frente de onda transmitida pelo sistema pode ser obtida por meio de medidas interferométricas.

Polinômios de Zernike são muito utilizados em óptica para a representação das aberrações, e são também ajustados à frente de onda extraída por interferômetros.

Para a análise dos sistemas teóricos, os softwares de projeto óptico possuem ferramentas que possibilitam avaliar a frente de onda de um sistema e extrair os coeficientes deste polinômio.

O objetivo deste trabalho é utilizar esses coeficientes para o estudo do estado de alinhamento de sistemas ópticos e extrair informações que auxiliem no refinamento do alinhamento óptico.

As redes neurais artificiais são empregadas para a análise destes coeficientes e estimativa dos erros de alinhamento.

4.3 Proposta de metodologia para alinhamento de sistemas ópticos

Metodologias para alinhamento de sistemas ópticos por meio das informações extraídas da frente de onda desses sistemas já vêm sendo propostas por diversos autores (6, 30, 32).

Uma metodologia similar, no que diz respeito à forma de se extrair informações de um sistema óptico que se deseja alinhar, é também utilizada neste trabalho. O que

diferencia esse método dos já existentes na literatura é a forma como os dados são analisados visando estimar o estado de alinhamento do sistema.

Redes neurais artificiais serão utilizadas para a estimativa dos erros de alinhamento em sistemas ópticos.

A descrição de uma metodologia completa será proposta neste trabalho, bem como uma versão dela que possibilita os estudos teóricos, foco desta pesquisa. Nesses estudos, os dados que a princípio são extraídos do sistema por métodos interferométricos, são gerados à partir de um modelo teórico utilizando o software de projeto óptico ZEMAX®.

Esses estudos permitem avaliar o desempenho de redes neurais artificiais como estimadores de erros de alinhamento, considerando: a escolha dos dados de entrada e seu pré-tratamento; diferentes tipos de erros de alinhamento: inclinação, descentralização e deslocamentos axiais; e efeitos de erros de forma nos elementos ópticos.

A metodologia completa consiste em seis etapas fundamentais:

- ETAPA 1: Modelagem do sistema óptico

  Consiste na elaboração do modelo teórico do sistema óptico a ser alinhado e, a partir do qual são extraídas as amostras de treinamento da rede.

- ETAPA 2: Extração das amostras de treinamento da RNA

  Consiste no processo de perturbação do modelo teórico do sistema e extração das amostras de treinamento e das amostras para avaliação do desempenho da rede (dados de entrada/saída).

- ETAPA 3: Elaboração da RNA

  Consiste em elaborar uma arquitetura/topologia de rede neural que aprenda a diferenciar os efeitos dos diferentes erros de alinhamento que podem estar presentes no sistema óptico sob análise e a estimá-los quantitativamente.

  A elaboração da rede abrange desde a escolha da arquitetura de rede a ser usada até os tipos de entradas e saídas apropriados para o seu treinamento.

- ETAPA 4: Treinamento da RNA

Consiste no treinamento da rede por meio da apresentação das amostras de treinamento e os ajustes dos seus diversos parâmetros de forma a obter a melhor configuração para o sistema que se deseja alinhar. Além disso, o treinamento deve ser verificado através da apresentação de amostras (entradas) que não foram apresentadas durante o processo de treinamento e que se tenha conhecimento das saídas corretas.

- ETAPA 5: Montagem experimental e extração de dados reais

  Consiste na montagem do *setup* experimental e na extração dos dados do sistema óptico (coeficientes dos polinômios de Zernike) que serão apresentados à RNA para que os erros de alinhamento presentes no sistema sejam estimados. No caso de uma análise puramente teórica, estes dados constituem dados extraídos do modelo teórico, da mesma forma que os dados de treinamento e de validação da rede. Na prática, estes dados são extraídos do sistema real e apresentados à rede para estimativa dos erros de alinhamento do sistema.

- ETAPA 6: Processo iterativo para alinhamento do sistema óptico

  Consiste na etapa iterativa do processo de alinhamento em que os dados reais são apresentados à rede, que estima os erros de alinhamento presentes fornecendo informações para realinhamento do sistema e nova extração de dados e assim sucessivamente até que os erros de posicionamento dos elementos no sistema estejam dentro das especificações para este sistema.

A Figura 29 apresenta um fluxograma geral da proposta de procedimento para alinhamento de sistemas ópticos por meio de medidas interferométricas.

No fluxograma da Figura 29, os dados extraídos do sistema que se deseja alinhar foram considerados como sendo medidas reais, obtidas por meio de medidas interferométricas do sistema. Contudo, nesta tese, para a verificação do desempenho da rede e a avaliação de seu desempenho, as amostras de teste, obtidas teoricamente da mesma forma que as amostras de treinamento, substituem essas medidas.

Basicamente, as etapas 1, 2, 3 e 4 consistem em modelar o sistema óptico, extrair as amostras de treinamento da rede e treiná-la. O ajuste da rede pode ser um

processo demorado uma vez que diversos parâmetros devem ser ajustados e o processo de treinamento pode ser longo. Estas são as etapas abordadas nesta tese.

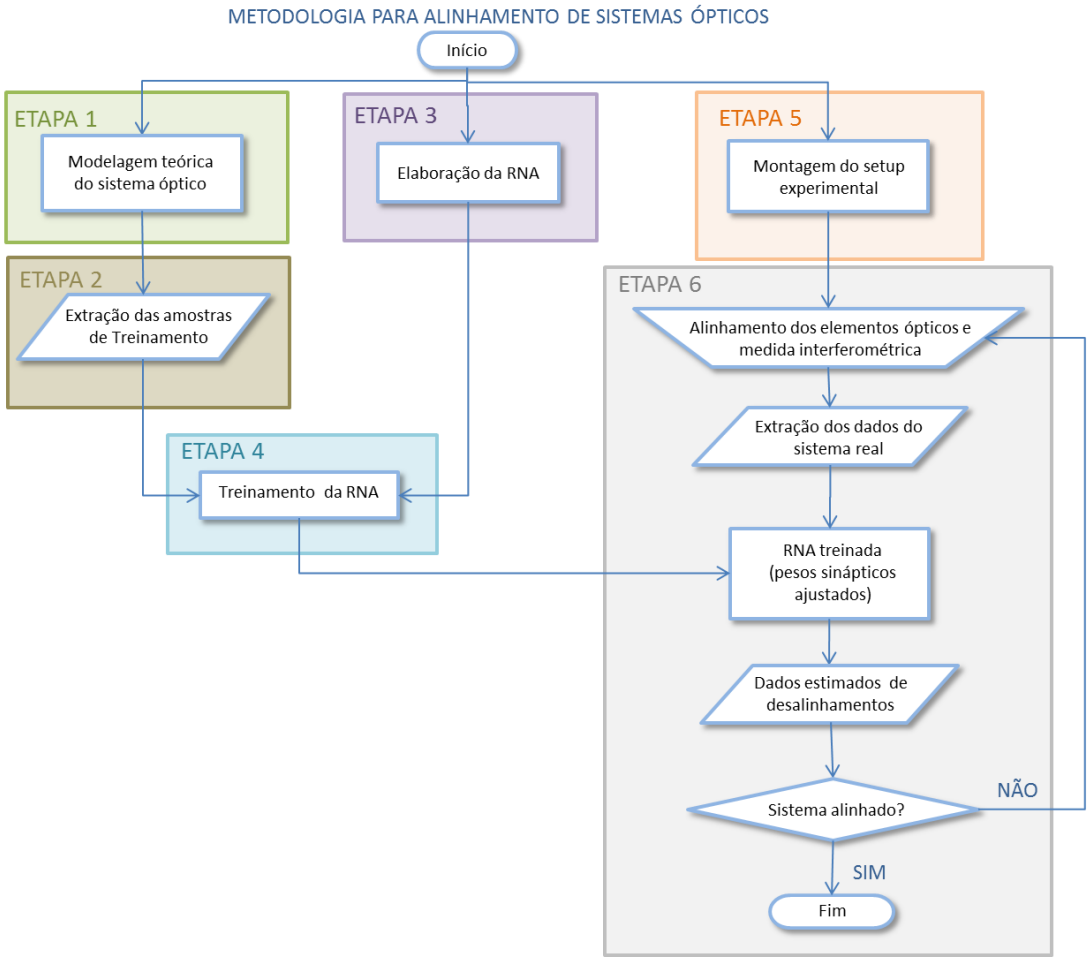

Figura 29 – Fluxograma do procedimento para alinhamento de sistemas ópticos por meio da análise dos polinômios de Zernike.

A etapa 5 é independente das demais, abrange a montagem experimental e a extração efetiva dos dados do sistema. O modelo teórico deve ser consistente com a montagem experimental, uma vez que a RNA deve ser treinada com dados provenientes do modelo, mas deverá prever o estado de alinhamento do sistema real. Desta forma, deve-se garantir a fidelidade entre o modelo e a montagem experimental.

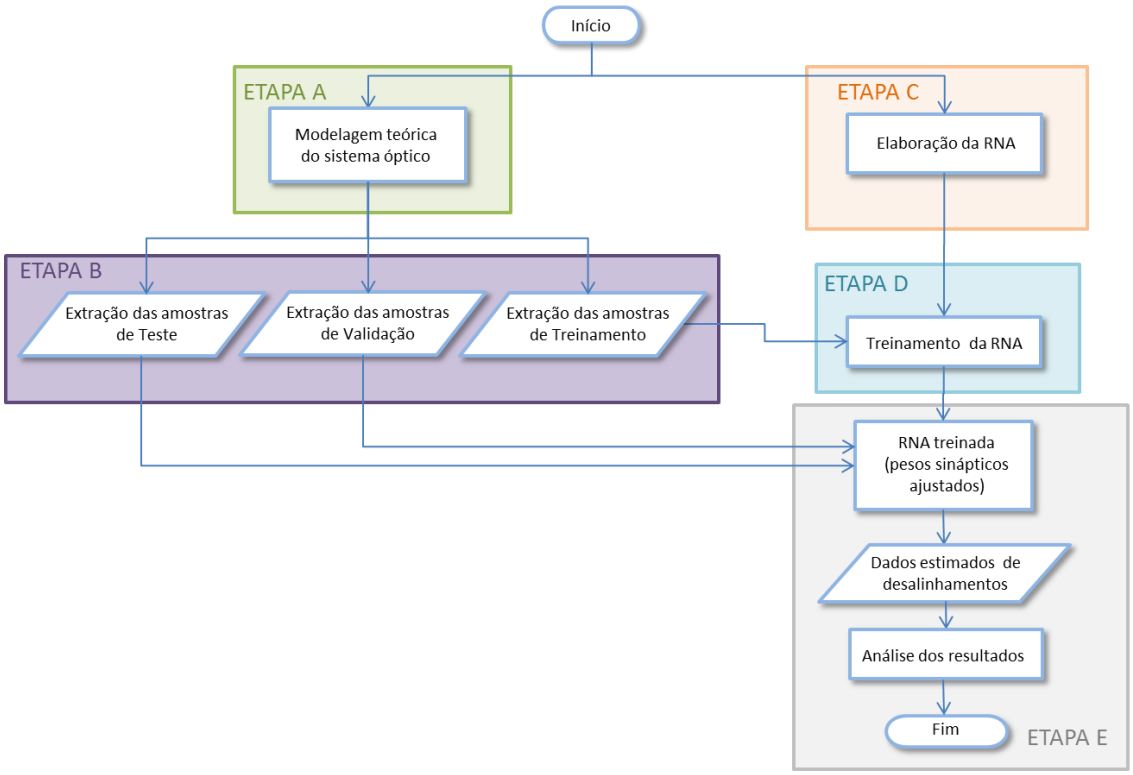

Figura 30 – Fluxograma do procedimento para alinhamento de sistemas ópticos por meio da análise dos polinômios de Zernike – análises teóricas.

A etapa 6 consiste no processo de alinhamento propriamente dito.

O processo de alinhamento pode ser longo, porque consiste em um processo iterativo, em que medidas interferométricas devem ser realizadas a cada etapa de alinhamento e reapresentadas à rede.

Cada uma das etapas será discutida detalhadamente a seguir.

A Figura 30 apresenta o fluxograma do processo de alinhamento utilizado neste trabalho, que é efetivamente realizado quando apenas dados teóricos estão disponíveis. Este fluxograma descreve as etapas usadas nas análises teóricas realizadas. A etapa A é equivalente à etapa 1 descrita anteriormente, a etapa B é equivalente às etapas 2 e 6, referentes à geração das amostras de treinamento e validação e da extração das amostras de teste, respectivamente. A etapa C é equivalente à etapa 3, a etapa D é equivalente à etapa 4 e a etapa E é equivalente à parte da etapa 6, onde os dados de teste (ou os dados reais) são apresentados à rede treinada e os valores dos erros de alinhamento são estimados e analisados.

Neste caso, não ocorre a parte iterativa do processo, que consiste no alinhamento propriamente dito.

Estes estudos buscam avaliar o desempenho de redes neurais artificiais quando aplicadas ao alinhamento de sistemas ópticos.

4.4 Modelagem do sistema óptico

O software de projeto óptico ZEMAX® foi utilizado para a formulação dos modelos teóricos.

Todos os sistemas analisados neste trabalho formam imagens de objetos situados a uma longa distância. Para estes sistemas o objeto está no infinito, ou seja, eles formam imagem quando o feixe incidente está colimado.

O modelo teórico do sistema, de onde são extraídos os dados de treinamento, deve ser consistente com a montagem experimental adotada para a medida da frente de onda transmitida pelo sistema, discutida na seção 4.8. Contudo, para a realização de análises puramente teóricas, um modelo simplificado do sistema é suficiente.

Como o objetivo deste trabalho é avaliar a sensibilidade e a eficiência das redes neurais artificiais quando usadas como estimadores de desalinhamento, modelos simplificados foram adotados. Nestes modelos, o sistema mantém a sua forma padrão, a mesma usada em processos de otimização. Contudo, quebras de coordenadas são inseridas no sistema por meio da implementação de superfícies do tipo *coordinate break* que possibilitam a inserção de perturbações (inclinação e/ou descentralização) no sistema (45).

4.5 Extração das amostras de treinamento

Neste trabalho, a proposta para a geração dos dados para análise do sistema, ou seja, das amostras de treinamento da rede neural, é que estas sejam extraídas do modelo teórico.

A extração dos dados de treinamento do modelo teórico é realizada como mostrado no diagrama da Figura 31.

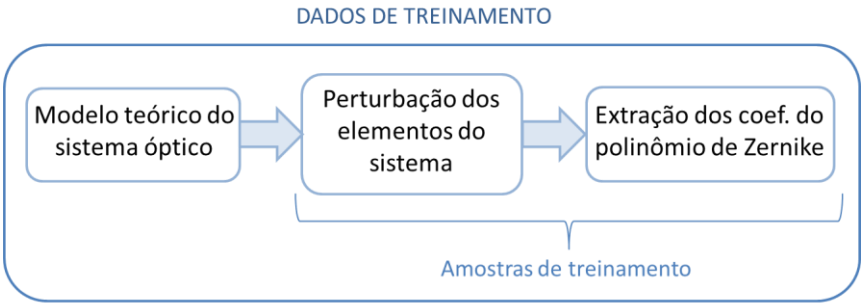

Figura 31 – Diagrama para a extração dos dados de treinamento a partir do modelo teórico do sistema.

Os dados de treinamento consistem em:

- ENTRADAS:
  Coeficientes dos polinômios de Zernike contendo as informações das aberrações presentes no sistema óptico em cada estado específico de desalinhamento do sistema.
- SAÍDAS:
  Consistem nos valores de desalinhamento(s) inserido(s) no sistema óptico (perturbações) correspondentes a cada conjunto de valores dos coeficientes de Zernike.

Os tipos de perturbações inseridas no sistema óptico são: inclinação, descentralização e deslocamento axial.

Além dos cinco graus de liberdade citados: inclinação no eixo *x* e no eixo *y*; descentralização no eixo *x* e no eixo *y* e deslocamento axial, há ainda a possibilidade de um sexto grau de liberdade que consiste na rotação do elemento em torno de seu eixo. Este movimento não causa nenhuma alteração no sistema se

não houver imperfeições nos elementos do sistema. Caso contrário, a rotação dos elementos pode alterar a frente de onda do sistema já que as deformações superficiais dos elementos, quando apresentam baixas frequências espaciais, podem ocasionar aberrações que podem, por sua vez, se somar ou se subtrair às demais.

4.5.1 Perturbação dos elementos do sistema e extração das amostras de treinamento

Por serem necessários muitos conjuntos de dados de entradas/saídas para a realização do treinamento da RNA, criou-se um programa para a realização automática do processo de extração desses dados.

O programa de projeto óptico ZEMAX® disponibiliza uma alternativa que possibilita que outros programas do Windows estabeleçam uma comunicação com ele e extraia dados ou mesmo modifique parâmetros em um sistema óptico. Essa comunicação é realizada por meio do *Dynamic Data Exchange* (DDE) que consiste em um protocolo definido dentro do sistema operacional do Windows para compartilhamento de dados entre os programas (45).

Esta comunicação entre dois programas atua da seguinte forma: um programa atua como sendo o “servidor” e o outro o “cliente”. O “cliente” requisita dados do “servidor” e este envia as informações para o “cliente”. No caso da comunicação com o ZEMAX, este é considerado o “servidor” e o outro programa o “cliente”. Assim, o ZEMAX pode ser acessado e usado como mecanismo para traçado de raios ou mesmo para cálculos específicos. Exemplos de softwares que permitem comunicação com o ZEMAX é o EXCEL® e o MATLAB®.

O código escrito para a extração dos dados de treinamento do ZEMAX foi criado utilizando-se o software MATLAB®.

Foi também utilizado um *toolbox* chamado MZDDE (71), que se encontra disponível livremente, sendo composto por uma série de funções pré-estabelecidas para acesso às funcionalidades do software ZEMAX utilizando a comunicação DDE, que

facilita significativamente a implementação de códigos que necessitam desta comunicação.

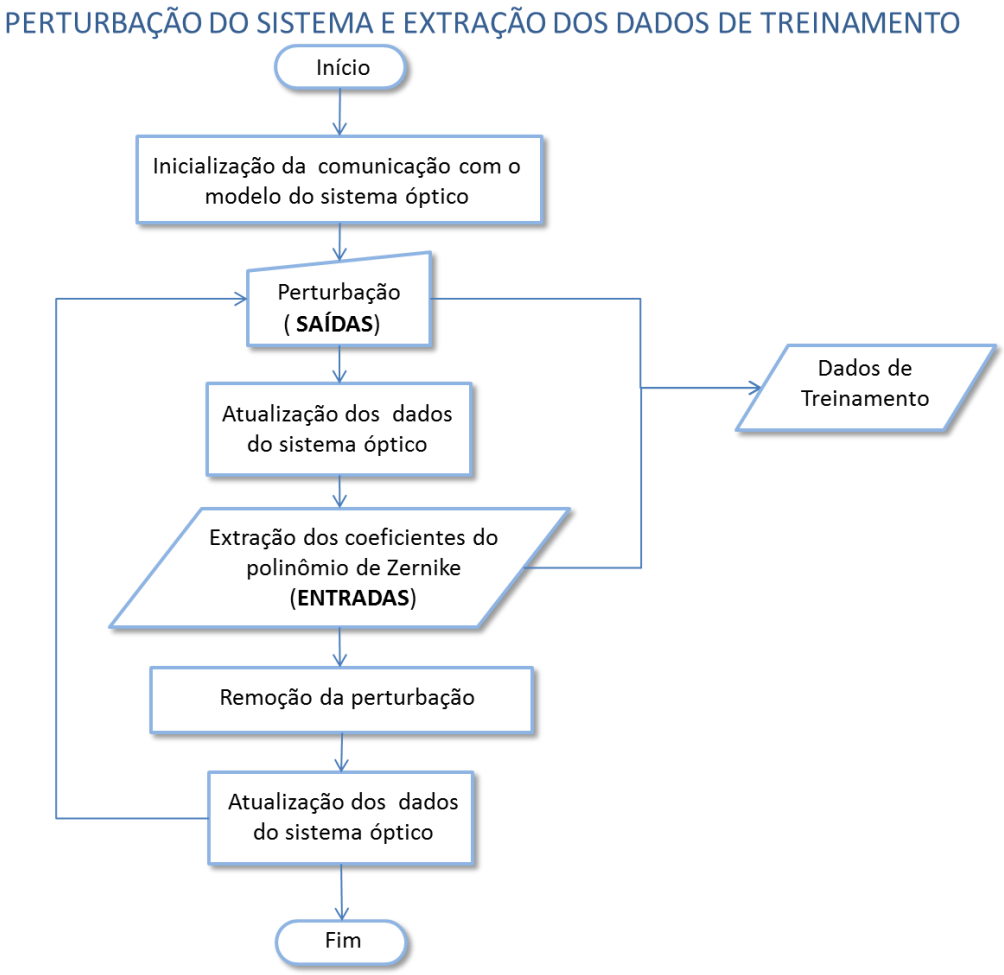

Figura 32 – Fluxograma para a extração dos dados de treinamento da RNA.

O código criado realiza modificações no sistema óptico desejado, inserindo a(s) perturbação(ões) no(s) elemento(s) do sistema, atualiza os dados do sistema, otimiza parâmetros do sistema, caso seja necessário, e extrai os valores dos polinômios de Zernike.

Por meio de iterações deste processo, podem-se criar inúmeros conjuntos de entradas/saídas para a composição dos dados de treinamento da RNA. As perturbações que se deseja inserir no sistema podem ser criadas individualmente para cada sistema. A Figura 32 apresenta um fluxograma do código criado.

Desse conjunto de amostras entradas/saídas, pode-se selecionar algumas para que sejam usadas como amostras de validação para a avaliação da RNA treinada.

4.5.2 Entradas

As entradas consistem em parte dos dados que compõem as amostras de treinamento. São os dados que serão apresentados à RNA e que carregam as informações extraídas do sistema usadas para a estimativa dos erros de alinhamento presentes no sistema óptico.

Os coeficientes dos polinômios de Zernike foram adotados como sendo as entradas a ser apresentadas à RNA.

O tipo de polinômio adotado foi o *Zernike Fringe Coefficients* (45).

Apenas alguns coeficientes deste polinômio são efetivamente relevantes para ser considerados dependendo do tipo de perturbação presente no sistema.

Conforme descrito no capítulo 2, aberração esférica não sofre alteração na presença de desalinhamentos dos tipos inclinação e descentralização (3); desta forma coeficientes relacionados a essa aberração não precisam ser considerados quando apenas esses dois tipos de erros de alinhamento estão presentes. Contudo o uso do coeficiente que corresponde a esta aberração e a *defocus* poderia ser uma alternativa para a estimativa de erros de posicionamento axial entre elementos.

Coeficientes relacionados às aberrações astigmatismo e coma são considerados por serem estas aberrações muito afetadas por descentralização e inclinação.

Distorção é afetada por erros de alinhamento e a presença deles faz com que surjam diversas novas dependências com o campo (3). Assim, os efeitos gerados por erros de alinhamento nessa aberração são importantes, porém, por serem mais complexos, podem causar confusão à rede.

É importante a realização de uma análise para avaliar, em cada sistema, quais os coeficientes que, necessariamente, devem ser considerados como dados de entrada para que a rede disponha de informações suficientes e consiga estimar os erros de alinhamento presentes no sistema. O uso de coeficientes adicionais, que não contribuem efetivamente para a melhora do desempenho do estimador, ocasiona um aumento dos dados e, consequentemente, um aumento do tempo de treinamento.

Os coeficientes que representam os termos de terceira ordem são os mais afetados por erros de alinhamento. Termos de ordens mais altas são também importantes porém não serão incluídos como dados de entrada, ao menos inicialmente, para reduzir a quantidade de dados apresentados à rede.

A Tabela 8 apresenta os diferentes coeficientes que podem compor os dados de entrada para a RNA. O conjunto de coeficientes usados pode ser alterado dependendo do sistema óptico a ser analisado e dos desalinhamentos considerados.

Tabela 8 – Polinômios de Zernike (*FRINGE*) que representam aberrações primárias que podem ser usados como dados de entrada para a RNA.

Zernike *Fringe* (ZEMAX)	**Aberração associada**
Z_2	Inclinação (*x*)
Z_3	Inclinação (*y*)
Z_4	*Defocus*
Z_5	Astigmatismo 3° ordem em *x*
Z_6	Astigmatismo 3° ordem em *y*
Z_7	Coma 3° ordem em *x*
Z_8	Coma 3° ordem em *y*
Z_9	Aberração esférica de 3° ordem

Fonte: MAHAJAN (4), WYANT (44), RADIANT ZEMAX (45), ZERNIKE Polynomials (49)

Conforme discutido por diversos autores na literatura, medidas da frente de onda realizadas apenas no eixo óptico podem não ser suficientes para garantir o alinhamento de um sistema óptico já que as novas dependências das aberrações com o campo podem levar ao aparecimento de mais de um ponto de aberração nula, conhecido como *node*. No caso do astigmatismo, os termos adicionais gerados por erros de alinhamento levam ao astigmatismo binodal, ou seja, dois pontos de astigmatismo nulo podem surgir no sistema.

Correções que buscam zerar esta aberração no eixo óptico podem, na verdade, estar apenas deslocando um dos pontos de zero para o eixo, dando a impressão de ter corrigido o sistema, porém este pode ainda estar desalinhado (20). Por outro lado, algumas metodologias de alinhamento óptico baseadas em medidas apenas sobre o eixo óptico também apresentaram bons resultados (70).

Como a abordagem proposta neste trabalho avalia diferentes aberrações, e não apenas astigmatismo, não podemos afirmar se medidas realizadas apenas no eixo

são ou não suficientes para a avaliação do sistema óptico. Por esta razão, uma análise mais detalhada foi realizada para se concluir se pontos de medição adicionais são ou não necessários para que a RNA consiga estimar os erros de alinhamento presentes no sistema óptico.

Assim, a literatura sugere que sejam consideradas medidas da frente de onda em cinco pontos diferentes ao longo do campo. Para um campo normalizado, os pontos a ser adotados para as medidas são: (0,0), (-1,0), (1,0), (0,-1), (0,1).

Nesta tese serão apresentados os resultados de uma análise realizada em um sistema reflexivo com dois elementos que visa avaliar quais coeficientes, dentre os apresentados na Tabela 8, são relevantes quando usados como dados de entrada para o treinamento de uma RNA, que visa estimar descentralização e inclinação, e de quais pontos ao longo do campo esses coeficientes precisam ser extraídos para que os dados de entrada contenham informações suficientes para que a rede tenha bom desempenho como estimador.

4.5.3 Saídas

As saídas da RNA consistem nos valores de desalinhamentos e na sua direção e sentido. A RNA será treinada para exibir, como resposta a cada conjunto de dados de entrada, os valores dos desvios de alinhamento presentes no sistema que correspondem às entradas apresentadas.

Quando apenas descentralização e inclinação são considerados, quatro (04) valores de saída são definidos para cada elemento do sistema referentes às quatro primeiras linhas da Tabela 9. Se o sistema apresentar apenas 1 elemento que sofre desalinhamentos, a RNA terá quatro saídas. Se o sistema apresentar dois elementos que podem estar desalinhados, a rede terá oito saídas e assim por diante.

Quando deslocamentos axiais também são considerados, cinco (05) valores de saída são definidos para cada elemento do sistema, como apresentado na Tabela 9.

Tabela 9 – Descrição das saídas da RNA, para cada elemento.

Saída	Nomenclatura	Tipo de desalinhamento
1	DX	Descentralização no eixo *x*
2	DY	Descentralização no eixo *y*
3	TX	Inclinação no eixo *x*
4	TY	Inclinação no eixo *y*
5	AX	Deslocamento axial

Neste caso, a RNA pode fornecer, como resposta, qual o tipo de erro de alinhamento presente, qual o elemento que apresenta o desalinhamento e a sua direção, sentido e magnitude.

Esta configuração de saídas também permite que combinações de erros de alinhamento sejam estimados ao mesmo tempo.

4.5.4 Amostras de validação e teste

As amostras de validação e teste são compostas por dados extraídos do modelo teórico, da mesma forma que as amostras de treinamento, que são apresentadas à RNA na busca dos valores de desvios de alinhamento.

Estas amostras constituem conjuntos de entradas/saídas que não são apresentados à rede durante o processo de treinamento. De modo geral são conjuntos com poucas amostras quando comparados ao conjunto de treinamento, porém em número suficiente para permitir uma análise do desempenho da rede.

As amostras de validação são apresentadas à rede para avaliar o seu desempenho, porém, para evitar um possível ajuste excessivo dos dados de validação, as amostras de teste são também apresentadas para testar o desempenho de generalização da rede, com um conjunto de dados diferente do de validação (15).

As saídas fornecidas pela rede para as amostras de validação e teste podem ser relacionadas com as saídas esperadas, ou seja, a informação de qual ou quais desalinhamentos estão presentes no sistema existe. Desta forma, é possível realizar uma avaliação do desempenho da rede, saber qual a porcentagem de acerto e quais os tipos de desalinhamentos estão causando confusão à rede.

Esse tipo de análise possibilita estimar a probabilidade de acertos que a rede terá quando os dados reais, efetivamente extraídos de um sistema real, forem apresentados a ela. Essa estimativa pode ficar mais realista se erros de forma, erros de medida e outros fatores perturbadores forem considerados.

Desta forma, os dados de validação e teste fornecem meios para avaliar o aprendizado da rede.

4.5.5 Tratamento dos dados

Os dados usados como entradas para a RNA e os dados de saídas, no caso dos dados de treinamento, extraídos do software de projeto óptico ou das medidas experimentais, não podem ser apresentados diretamente à rede.

Um tratamento destes dados e a sua preparação antes de serem apresentados à rede são de fundamental importância no processo de treinamento da RNA para a estimativa dos erros de alinhamento de um sistema óptico. A seguir são apresentados alguns tratamentos realizados nos dados extraídos do sistema teórico.

- DADOS DIFERENCIAIS

  Os dados extraídos do modelo teórico constituem os valores dos coeficientes dos polinômios de Zernike em cada estado de alinhamento do sistema óptico. Estes dados, da forma como são extraídos, não fornecem de fato uma informação clara à rede. Isso porque eles carregam também informações do estado alinhado do sistema.

  Desta forma, se os valores dos coeficientes extraídos do sistema, quando este se encontra no estado alinhado, for descontado dos valores deste mesmo coeficiente obtidos quando há elementos desalinhados no sistema, o valor final representa o desvio do valor deste coeficiente em relação ao seu valor “ideal” que é o valor para o sistema em seu estado perfeitamente alinhado.

  Isso significa que uma informação mais clara do estado de alinhamento do sistema é fornecida quando apenas as diferenças dos valores dos coeficientes são apresentadas à rede.

Por esse motivo, os dados que formam as entradas apresentadas à rede não são diretamente os coeficientes dos polinômios de Zernike, mas dados diferenciais que quantificam a mudança do valor deste polinômio relativamente ao seu valor nominal correspondente ao sistema em seu estado alinhado.

- DADOS NORMALIZADOS

São encontrados na literatura diversos relatos de que o tratamento adequado dos dados de entrada resulta em um melhor desempenho da RNA.

A normalização dos dados de entrada da RNA é muito importante pois a função de ativação tende a ficar com valores saturados fora da faixa de valores de -1 a 1. Ela é realizada entre os valores de um mesmo dado de entrada, considerando todas as amostras que serão apresentadas à rede. Da mesma forma, as saídas também são normalizadas.

A “normalização” dos dados consiste em uma transformação linear dos dados originais, e nem sempre significa uma normalização propriamente dita (72).

Esses tratamentos dos dados devem ser efetuados antes que as informações sejam apresentadas à rede e facilitam o seu entendimento e aprendizado.

4.6 Elaboração da RNA

Como descrito no Capitulo 2, redes neurais artificiais constituem uma ferramenta com grande potencial na aplicação a processos de alinhamento fino de sistemas ópticos devido à sua capacidade de aprendizagem e generalização. Além disso, apresentam bons resultados quando se trata de problemas com não linearidades.

A rede denominada Perceptron de Múltiplas Camadas (MLP, *Multilayer Perceptron*) foi adotada para essa tarefa por ser a mais apropriada porque faz combinação linear de funções não lineares. Ela é utilizada na resolução dos mais variados problemas. Segundo Haykin (15), o MLP treinado com o algoritmo de retropropagação pode ser visto como um veículo prático para a realização de um mapeamento não linear de entrada/saída de natureza geral. Ainda segundo o mesmo autor, o Teorema da Aproximação Universal afirma que uma única camada escondida é suficiente para

um MLP computar uma aproximação uniforme para um dado conjunto de treinamento formado por um conjunto de entradas/saídas. No entanto, esse teorema não afirma que uma única camada escondida é “ótima” no que diz respeito a tempo de aprendizagem, facilidade de implementação ou generalização (15).

Esse teorema considera que a única camada escondida possui tamanho ilimitado e que a função contínua a ser aproximada é dada, o que não é verdade na maioria das aplicações práticas. MLPs que possuem uma única camada escondida têm o problema que os neurônios nesta camada tendem a interagir entre si globalmente podendo ocasionar dificuldades de aproximação em um ponto, sem piorá-la em outro. Para duas camadas escondidas, porém, o processo de aproximação fica mais gerenciável e cada camada assume tarefas específicas, ficando a primeira encarregada das características locais e a segunda das características globais (15). Maiores informações a respeito desse assunto são encontradas na referência citada.

4.7 Treinamento e avaliação do desempenho da RNA

4.7.1 Treinamento da RNA

Os parâmetros da RNA são ajustados por meio da repetição do processo de treinamento e posterior análise dos resultados.

A avaliação do desempenho da rede é realizada por meio de amostras de validação e do conjunto de teste. No conjunto de dados extraídos do modelo teórico, algumas amostras são separadas das que irão compor as amostras de treinamento e serão usadas para a validação da rede.

Cerca de 15% das amostras (conjunto entrada/saída) extraídas do modelo compõem as amostras de validação e outros 15% compõem o conjunto de teste.

Esses conjuntos não são apresentados à rede durante o processo de treinamento.

As amostras extraídas do modelo são embaralhadas de forma aleatória, após passar pelo pré-tratamento, e são divididas nestes três conjuntos (treinamento, validação e teste) também de forma aleatória.

A rede será treinada repetidas vezes, sempre inicializando os pesos e *bias* de forma randômica, e separando estes três conjuntos de dados de forma aleatória.

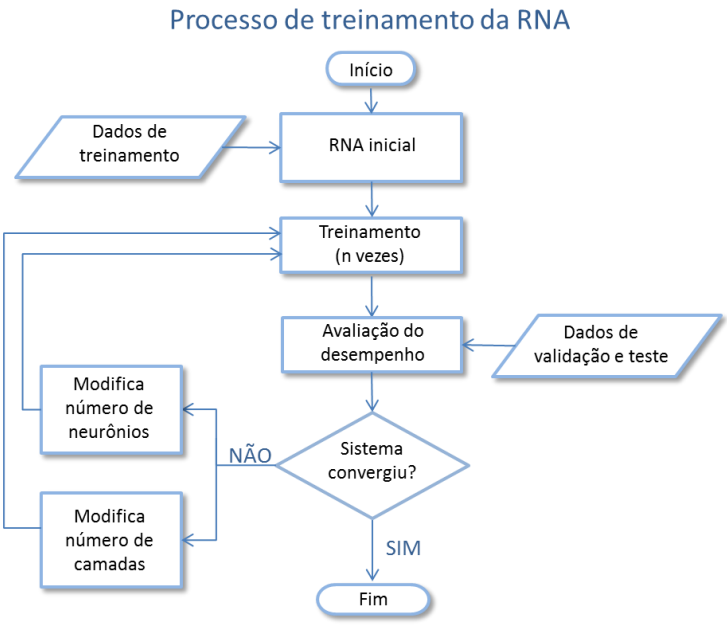

Figura 33 – Fluxograma do processo de treinamento das RNAs.

Após vários treinamentos seguidos de validação e teste, os resultados serão analisados para averiguação do desempenho. A Figura 33 apresenta, resumidamente, o processo adotado para o treinamento das redes neurais artificiais.

4.7.2 Avaliação do desempenho da RNA

Para a avaliação do desempenho da rede para cada par saída esperada (SE)/saída obtida (SO), a diferença entre esses valores (SE-SO) é calculada e apresentada em forma de histogramas de erros.

Além disso, um gráfico desta diferença em função do número da amostra é apresentado para avaliação do desempenho da rede. Esse gráfico possibilita avaliar a distribuição dos erros entre as amostras.

O valor do EQM para os dados de teste nos dá uma boa indicação do desempenho da rede. Além disso, o gráfico de EQM em função do número de épocas é apresentado para as amostras de treinamento, validação e teste como forma de avaliar o desempenho da rede durante o treinamento.

4.8 Montagem experimental e extração de dados reais

A aplicação da metodologia de alinhamento proposta depende da medida das aberrações de frente de onda do sistema que se pretende alinhar.

As aberrações do sistema, que carregam as informações necessárias para a estimativa do estado de alinhamento do sistema, são analisadas a partir dos coeficientes dos polinômios de Zernike.

O aparato experimental deve permitir a realização de medidas interferométricas do sistema em diferentes posições ao longo do campo, seguindo o que foi adotado no modelo teórico para a extração das amostras de treinamento.

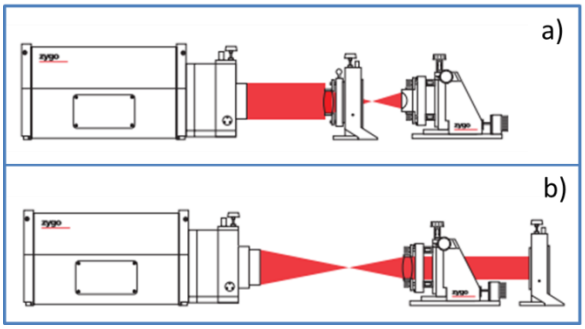

Figura 34 – Montagens experimentais possíveis para os sistemas analisados nesta tese, que formam imagens de objetos localizados no infinito; a) O sistema é posicionado de forma que um feixe colimado incida sobre ele e uma esfera de referência, côncava ou convexa, é usada para refletir o feixe de volta ao interferômetro; b) O sistema é posicionado de forma que o seu ponto de foco coincida com o ponto de foco do feixe do interferômetro; o feixe emergente é então colimado e um espelho plano de referência é usado para retrorrefletir o feixe. Fonte: ZYGO CORPORATION (74).

Esses dados constituem valores reais, extraídos, por exemplo, por meio de medidas interferométricas do sistema e que efetivamente se deseja processar para que sejam encontrados os elementos desalinhados, os tipos de desalinhamentos e as suas magnitudes.

Considerando as características dos sistemas analisados neste trabalho e o interferômetro disponível na empresa Opto Eletrônica (ZYGO GPI/XP) para a realização de medidas experimentais, duas montagens são possíveis (74), conforme apresentado na Figura 34.

4.9 Processo iterativo para alinhamento de sistemas ópticos

Os coeficientes de Zernike, extraídos a partir da frente de onda do sistema, obtida da primeira medida interferométrica do sistema real, são tratados e apresentados à rede treinada, que estimará os erros de alinhamento.

Estes valores de desalinhamentos estimados devem ser compensados no sistema real, por meio do reajuste do posicionamento do elemento desalinhado, e novas medidas interferométricas devem ser realizadas e apresentadas à rede.

Este processo deve ser repetido até que o desempenho do sistema seja satisfatório.

Mais uma vez é importante ressaltar que o foco deste trabalho consiste apenas no estudo teórico da metodologia proposta, referente às etapas de 1 a 4, por isso as montagens experimentais e o processo iterativo de alinhamento são citados brevemente.

Trabalhos futuros serão realizados buscando estudar e desenvolver detalhadamente estas duas etapas da metodologia para alinhamento óptico proposta nesta tese.

5 RESULTADOS E DISCUSSÕES

5.1 Introdução

Neste capítulo são apresentados os resultados obtidos da aplicação da metodologia proposta no Capítulo 4 nos seguintes sistemas:

- Sistema reflexivo com dois elementos;
- Câmera multiespectral para sensoriamento remoto orbital.

O sistema reflexivo com dois elementos, do tipo *Cassegrain* clássico, foi analisado no Capítulo 3 quanto a sua sensibilidade ao alinhamento quando inclinações e descentralizações estão presentes no espelho secundário. Neste capítulo, este sistema é analisado visando avaliar o desempenho de redes neurais artificiais quando usadas como estimadores de desalinhamentos. São analisados inclinações, descentralizações e deslocamentos axiais. Além disso, são abordados os efeitos que erros de forma ocasionam no desempenho do estimador.

A metodologia proposta é também aplicada, teoricamente, ao sistema óptico da câmera multiespectral para sensoriamento remoto orbital MUX, discutido brevemente na seção 1.2. O objetivo dessa análise é mostrar que essa metodologia de alinhamento apresenta bons resultados quando aplicada a sistemas ópticos refrativos complexos.

5.2 Sistema reflexivo com perturbações no espelho secundário

Esta seção apresenta um exemplo da aplicação do modelo proposto utilizando para isso dados extraídos do sistema teórico ideal.

O sistema escolhido para este exemplo consiste em um telescópio reflexivo rotacionalmente simétrico, composto por dois elementos, do tipo *classical Cassegrain*. Esse sistema foi escolhido por possuir apenas duas superfícies. A superfície do primeiro espelho é a referência do sistema e, portanto, apenas uma superfície, a do espelho secundário, pode sofrer desalinhamentos.

Os elementos do sistema são considerados perfeitos, ou seja, não possuem deformações superficiais, que podem ser ocasionadas durante o processo de fabricação, ou pelos suportes e estruturas mecânicas.

A modelagem adotada é simplificada já que não são extraídos dados reais desse sistema.

O objetivo é mostrar que a metodologia proposta para o alinhamento do sistema, utilizando redes neurais artificiais como estimador, é capaz de identificar os desvios de alinhamento presentes no sistema e de prever, com base nos coeficientes dos polinômios de Zernike extraídos do sistema, quais desalinhamentos estão presentes no espelho secundário.

É apresentada uma comparação do desempenho da rede diante de diferentes estratégias adotadas no seu treinamento como, por exemplo, a escolha de diferentes conjuntos de coeficientes de Zernike como dados de entrada e a influência do número de posições no campo onde esses coeficientes são analisados.

5.2.1 Modelagem do sistema óptico

O modelo do sistema óptico adotado nesta análise simplificada baseia-se no modelo padrão do sistema, onde as superfícies são posicionadas sequencialmente e o feixe de raios cruza os elementos apenas uma vez, incidindo, primeiramente, no espelho primário e, em seguida, no espelho secundário. A Figura 35 apresenta um desenho esquemático do modelo adotado.

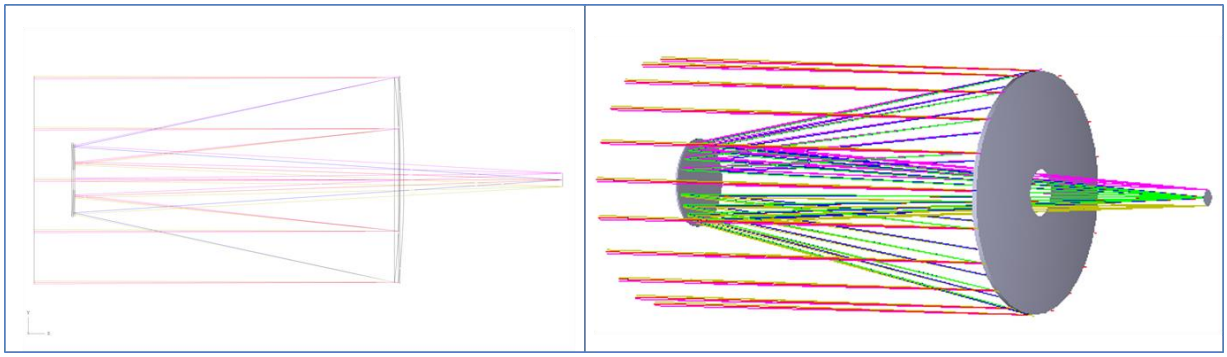

Figura 35 – Modelo de um telescópio do tipo *classical Cassegrain*.

A Tabela 10 apresenta as características do sistema óptico analisado.

Tabela 10 – Parâmetros do sistema óptico.

Parâmetro	Espelho primário	Espelho secundário	Sistema
Distância focal efetiva – EFL (mm)	-	-	3.750
Distância focal posterior – BFL (mm)	-	-	1.200
Separação entre os espelhos (mm)	-	-	-800
Raio de curvatura (mm)	-2.352,941	-1.097,143	-
Constante cônica (k)	-1,00	-3,664	-
Termo asférico de 4° ordem	-4,901E-16	-	-
Termo asférico de 6° ordem	5,872E-24	-	-
Termo asférico de 8° ordem	3,974E-29	-	-

As amostras de treinamento foram extraídas mediante a inserção das perturbações apresentadas na Tabela 11 no segundo elemento, já que o primeiro é considerado a referência do sistema. O intervalo de valores possíveis de desalinhamentos também é apresentado na tabela, assim como a amostragem adotada. Apesar de haver 16 combinações possíveis de saídas, já que a rede possui 4 saídas, respectivamente DX, DY, TX e TY, apenas 8 tipos foram considerados nesta análise: Alinhado, DX, DY, TX, TY, DX e DY, DX e TX, e TX e TY.

Por se tratar de um sistema cuja análise é apenas teórica, o intervalo de valores de desalinhamentos considerado, assim como a amostragem foram escolhidos arbitrariamente. Em uma aplicação real, estes valores devem ser definidos de acordo com as tolerâncias do sistema, o projeto mecânico e os ajustes disponíveis.

Tabela 11 – Descrição dos intervalos de desalinhamentos e amostragens adotadas.

Desalinhamentos	Intervalo	Amostragem
Descentralização em *x*	-1,0 a 1,0 mm	0,01 mm
Descentralização em *y*	-1,0 a 1,0 mm	0,01 mm
Inclinação em *x*	-1,0 a 1,0°	0,01 °
Inclinação em *y*	-1,0 a 1,0°	0,01 °
Descentralização em *x* e *y*	-1,0 a 1,0 mm	0,01 mm
Inclinação em *x* e *y*	-1,0 a 1,0°	0,01 °
Descentralização em *x* e Inclinação em *x*	-1,0 a 1,0 mm e -1,0 a 1,0°	0,01 mm e 0,01 °

5.2.2 Efeitos da escolha do conjunto de dados de entrada

Esta seção apresenta comparações entre diferentes estratégias de treinamento da rede que visam avaliar o seu desempenho mediante a apresentação de diferentes conjuntos de dados de entrada.

Para essa análise foi adotada uma rede com duas camadas escondidas, que, de acordo com alguns autores (15) é suficiente, na maioria das aplicações. O número de neurônios nas camadas e o tipo de função de ativação adotados em cada estratégia são apresentados e correspondem aos valores que resultaram no melhor desempenho da rede, obtidos entre diferentes treinamentos com diferente número de neurônios em cada camada escondida. Tais valores foram escolhidos arbitrariamente visando encontrar a faixa de número de neurônios que oferecesse o melhor desempenho. A camada de saída apresenta 4 neurônios representando cada um dos desalinhamentos na única superfície que sofre perturbações (espelho secundário).

5.2.2.1 Estratégias adotadas

Quatro estratégias diferentes foram adotadas para a seleção dos coeficientes de Zernike a ser usados como dados de entrada para o treinamento da rede. A Tabela 12 apresenta as possibilidades analisadas. São apresentados os coeficientes analisados e as posições no campo de onde foram extraídos. A escolha destes coeficientes tem como base as teorias de aberrações de onda, apresentadas no capítulo 2, que predizem os efeitos dos desalinhamentos nas aberrações de sistemas ópticos.

A estratégia I consiste em um conjunto de dados de entrada com apenas 4 elementos correspondentes às aberrações de terceira ordem: coma e astigmatismo, extraídas apenas no eixo óptico do sistema.

Tabela 12 – Estratégias adotadas para o conjunto de dados de entrada para treinamento da rede.

Estratégia I	Estratégia II	Estratégia III	Estratégia IV
$Z_5[(0,0)]$	Z_2 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	Z_5 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	$Z_2[(0,0]$
$Z_6[(0,0)]$	Z_3 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	Z_6 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	$Z_3[(0,0)]$
$Z_7[(0,0)]$	Z_4 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	Z_7 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	$Z_4[(0,0)]$
$Z_8[(0,0)]$	Z_5 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	Z_8 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$	$Z_5[(0,0)]$
	Z_6 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$		$Z_6[(0,0)]$
	Z_7 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$		$Z_7[(0,0)]$
	Z_8 $[(0,0),(-1,0),(1,0),(0,-1),(0,1)]$		$Z_8[(0,0)]$

A estratégia II consiste em um conjunto com 35 elementos sendo 7 coeficientes extraídos em 5 posições no campo. Adicionalmente, aos coeficientes escolhidos na estratégia I, foram considerados coeficientes relacionados à inclinação nos eixos *x* e *y* e a *defocus*.

A estratégia III considera um conjunto com 20 dados de entrada composto por coeficientes relacionados às aberrações coma e astigmatismo, porém extraídos em 5 diferentes posições no campo.

Por fim, a estratégia IV considera 7 coeficientes, os mesmos da estratégia II, porém extraídos apenas no ponto central do campo, sobre o eixo óptico.

5.2.3 Resultados e discussões

A seguir são apresentados os resultados das análises de desempenho da rede mediante a apresentação dos diferentes conjuntos de dados de entrada. Esta análise baseia-se na avaliação das saídas da rede mediante amostras que não foram apresentadas a ela previamente durante a fase de treinamento. O total de amostras avaliadas foi de 18.301 e correspondem a 15% do total de amostras (122.007) extraídas do modelo teórico. Outros 15% foram usados para a validação do treinamento e 70% dessas amostras foram efetivamente usadas para treinar a rede. A cada treinamento, os conjuntos de amostras de treinamento, validação e teste foram separados de forma aleatória, não sendo sempre as mesmas amostras a ser consideradas nas análises apresentadas a seguir.

A topologia de rede usada consiste em um MLP com duas camadas escondidas, cujas funções de ativação usadas foram a logística ou a função tangente hiperbólica, e a camada de saída, com função de ativação linear.

As funções de ativação usadas apresentaram resultados similares nos casos analisados. Em cada estratégia, serão apresentados os resultados da topologia que apresentou melhor desempenho.

Estratégia I:

Os resultados apresentados a seguir correspondem a uma rede com 32 e 16 neurônios na primeira e segunda camadas ocultas, respectivamente, e 4 neurônios na camada de saída. A função de ativação usada nesta estratégia, que resultou no melhor desempenho, foi a função logística, no caso dos neurônios das camadas escondidas, e função linear para a camada de saída.

A Tabela 13 apresenta o EQM para cada uma das 4 saídas e o valor total diante das amostras de teste. O valor médio para cada saída e os desvios dos valores do EQM total em cada treinamento em relação ao seu valor médio também são apresentados. A Tabela 14 apresenta a porcentagem das amostras com erro, em módulo, inferior a 0,005mm ou 0,005°. A Figura 36 apresenta o EQM em função do número de épocas de treinamento para o treinamento 1. O valor mínimo atingido durante treinamento foi 3,63E-6. Os gráficos apresentados na Figura 37 e na Figura 38 correspondem ao treinamento 1. Nos gráficos, o erro para cada saída foi calculado como sendo $Erro = d_j(k) - y_j^s(k)$, onde $d_j(k)$ é o valor esperado e $y_j^s(k)$ é o valor obtido pela rede para a mesma amostra.

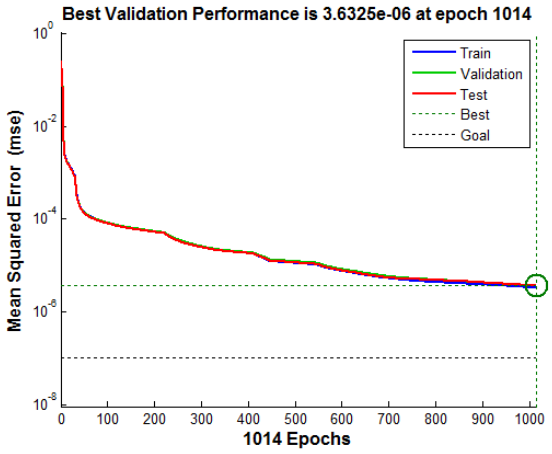

Figura 36 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia I, treinamento 1.

Tabela 13 – EQM para cada uma das quatro saídas da RNA e valor total, média dos valores de cinco treinamentos e desvio dos valores de EQM total em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia I – amostras de teste.

ESTRATÉGIA I						
Treinamento	**DX**	**DY**	**EQM TX**	**TY**	**TOTAL**	**Desvio (EQM TOTAL)**
Conj. 1	2,55E-5	3,08E-5	1,20E-6	9,29E-7	1,46E-5	-1,16E-5
Conj. 2	6,64E-5	8,61E-5	4,13E-6	5,49E-6	4,05E-5	1,43E-5
Conj. 3	2,87E-5	3,91E-5	1,58E-6	8,01E-7	1,75E-5	-8,66E-6
Conj. 4	5,44E-5	9,14E-5	3,56E-6	5,48E-6	3,87E-5	1,25E-5
Conj. 5	3,36E-5	4,11E-5	1,95E-6	1,16E-6	1,95e-5	-6,66E-6
Valor Médio	4,17E-5	5,77E-5	2,48E-6	2,77E-6	2,62E-5	

Tabela 14 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das quatro saídas da RNA e valor total, treinamentos com as amostras referentes à estratégia I – amostras de teste.

ESTRATÉGIA I					
Treinamento	**DX**	**DY**	**TX**	**TY**	**TOTAL**
Conj. 1	89,8	88,3	99,6	99,8	94,4
Conj. 2	71,4	62,7	97,5	95,5	81,8
Conj. 3	87,6	83,4	99,3	99,8	92,5
Conj. 4	74,6	67,5	98,0	95,4	83,9
Conj. 5	80,7	77,8	99,2	99,6	89,3
Valor médio	80,8	75,9	98,7	98,0	88,4

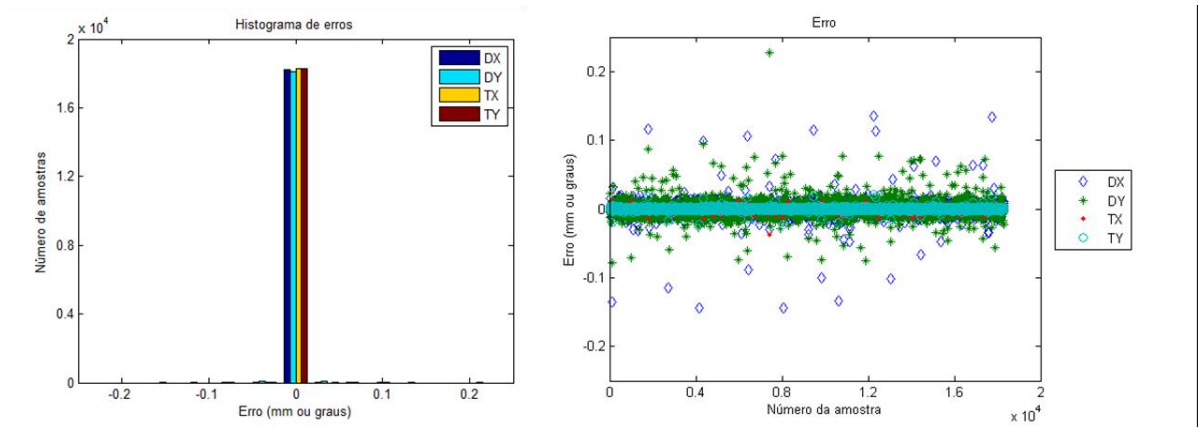

Figura 37 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia I, treinamento 1.

O EQM em função do número de épocas de treinamento indica boa convergência da rede. A rede treinada com amostras da estratégia I, onde apenas os coeficientes de Z_5 a Z_8 extraídos apenas no eixo óptico foram usados para seu treinamento, apresenta bom desempenho quando usada como estimador de descentralização e inclinação.

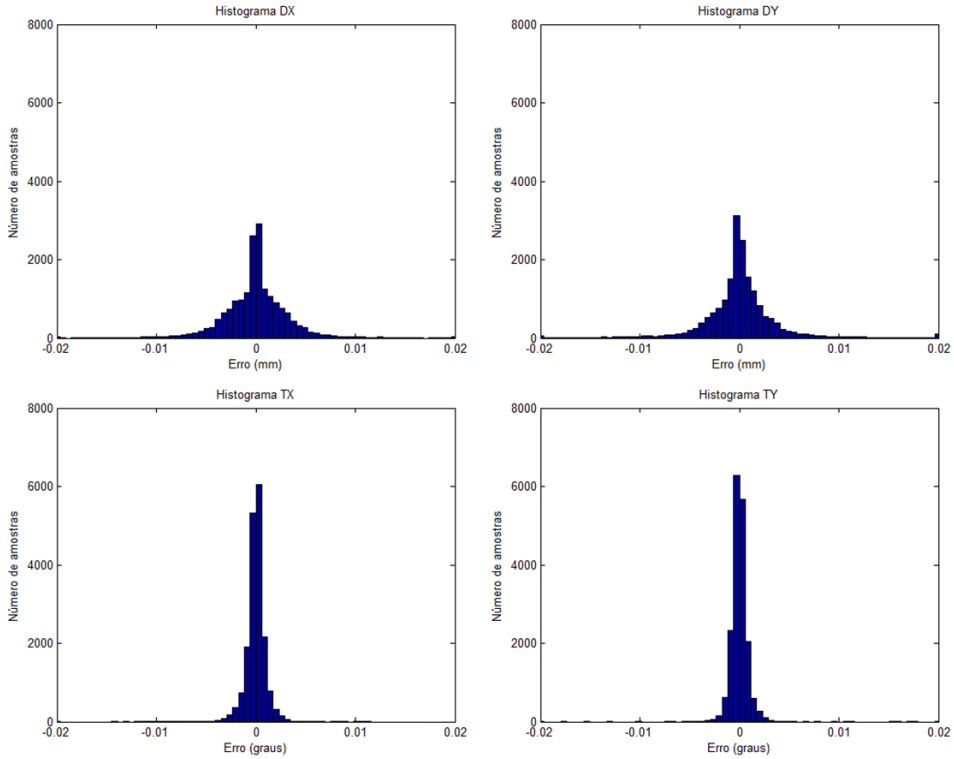

Figura 38 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia I, treinamento 1.

Estratégia II:

Os resultados apresentados a seguir correspondem a uma rede com 8, 16 e 4 neurônios respectivamente. As camadas escondidas possuem função de ativação do tipo tangente hiperbólica e a camada de saída possui função de ativação linear.

A Tabela 15 apresenta o EQM para cada uma das 4 saídas e o EQM total em cada treinamento diante das amostras de teste. Esta tabela apresenta também o valor médio de cada saída e os desvios do EQM total em cada treinamento em relação ao seu valor médio. Para esta estratégia a porcentagem das amostras com erro, em módulo, inferior a 0,005mm ou 0,005° para cada uma das 4 saídas e o valor total em cada treinamento, diante das amostras de teste, foi de 100%. A Figura 39 apresenta o EQM em função do número de épocas de treinamento para o treinamento 1. O valor mínimo atingido durante treinamento foi 6,55E-9. A Figura 40 apresenta o histograma de erros e um gráfico de erro em função do número da amostra para as quatro saídas da rede para o treinamento 1. A Figura 41 apresenta os histogramas de erros para cada uma das saídas, treinamento 1.

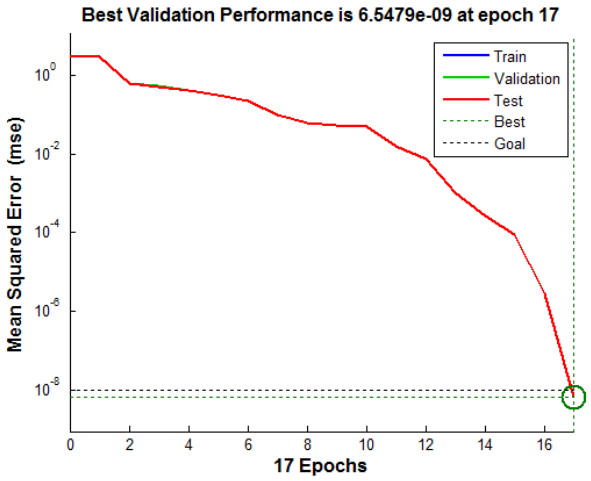

Figura 39 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia II, treinamento 1.

Tabela 15 – EQM para cada uma das quatro saídas da RNA e EQM total, valores médios e desvio do EQM total em cada treinamento em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia II – amostras de teste.

ESTRATÉGIA II						
Treinamento	**DX**	**DY**	**EQM TX**	**TY**	**TOTAL**	**Desvio (EQM TOTAL)**
Conj. 1	1,89E-8	1,62E-9	4,12E-9	1,52E-9	6,54E-9	-2,27E-9
Conj. 2	1,61e-8	3,61e-9	7,18e-9	1,06E-8	9,36E-9	5,54E-10
Conj. 3	4,11E-9	4,19E-9	1,12E-8	1,43E-8	8,43E-9	-3,76E-10
Conj. 4	1,29E-8	9,95E-9	9,89E-9	6,51E-9	9,80E-9	9,94E-10
Conj. 5	7,63E-9	1,16E-8	1,04E-8	9,94E-9	9,90E-9	1,09E-9
Valor Médio	1,19E-8	6,19E-9	8,56E-9	8,57E-9	8,81E-9	

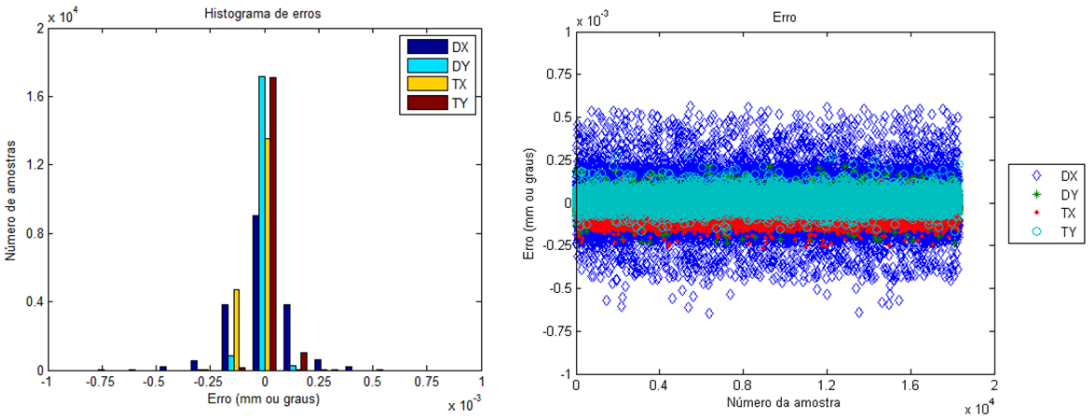

Figura 40 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia II, treinamento 1.

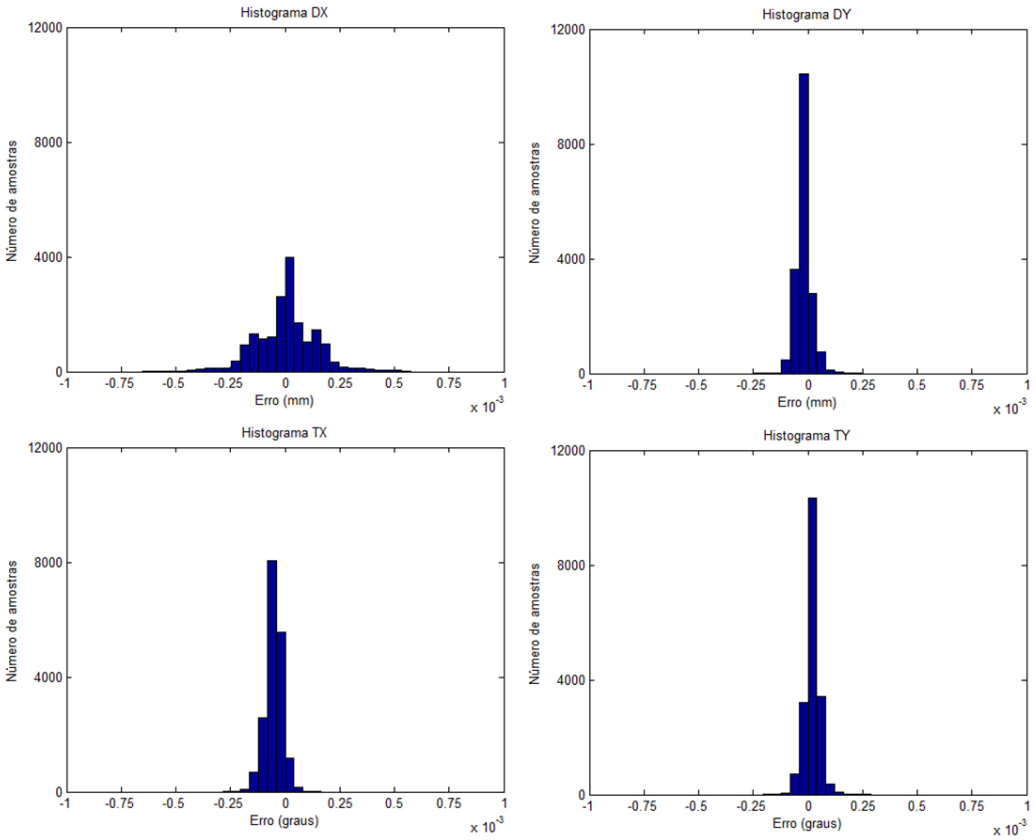

Figura 41 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia II, treinamento 1.

Os histogramas e gráficos de erro mostram que o erro entre o valor esperado e o obtido para as quatro saídas ficou abaixo de 0,001 mm ou 0,001°, o que indica bom desempenho da rede quando treinada com amostras referentes à estratégia II.

Estratégia III:

Os resultados apresentados a seguir correspondem a uma rede com 8, 16 e 4 neurônios respectivamente. As camadas escondidas possuem função de ativação do tipo tangente hiperbólica e a camada de saída possui função de ativação linear.

A Tabela 16 apresenta o EQM para cada uma das 4 saídas e o EQM total em cada treinamento diante das amostras de teste. Esta tabela apresenta também o valor médio de cada saída e os desvios do EQM total em cada treinamento em relação ao seu valor médio. Para esta estratégia novamente a porcentagem das amostras com erro, em módulo, inferior a 0,005mm ou 0,005° para cada uma das 4 saídas e o valor total em cada treinamento, diante das amostras de teste, foi de 100%. A Figura 42 apresenta o EQM em função do número de épocas de treinamento. O valor mínimo

atingido durante o treinamento 1 foi 8,5E-9. A Figura 43 apresenta o histograma de erros e o gráfico de erro em função do número da amostra para as quatro saídas da rede no caso do treinamento 1. A Figura 44 apresenta os histogramas de erros para cada uma das quatro saídas, também para o treinamento 1. Estes gráficos mostram que o erro ficou abaixo de 0,001 mm ou 0,001 °.

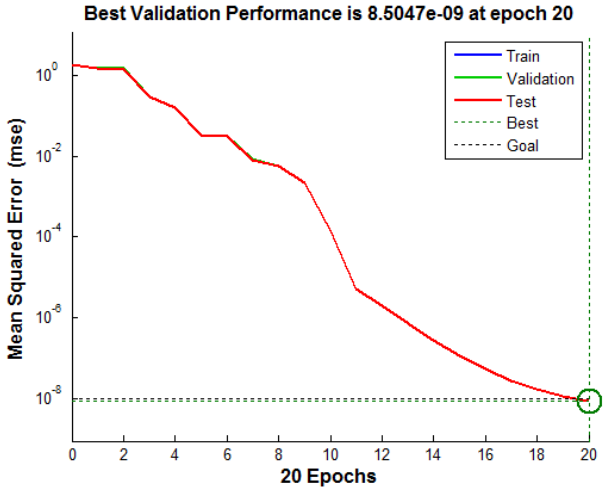

Figura 42 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia III, treinamento 1.

Tabela 16 – EQM para cada uma das quatro saídas da RNA e valor total, valor médio e desvio do EQM total em cada treinamento em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia III – amostras de teste.

ESTRATÉGIA III						
Treinamento	**DX**	**DY**	**EQM TX**	**TY**	**TOTAL**	**Desvio (EQM TOTAL)**
Conj. 1	1,21E-8	7,00E-9	4,71E-9	1,04E-8	8,56E-9	2.04E-10
Conj. 2	9,04E-9	2,87E-9	9,38E-9	1,34E-8	8,69E-9	3,34E-10
Conj. 3	4,76E-9	8,35E-9	3,87E-9	1,40E-8	7,75E-9	-6,06E-10
Conj. 4	1,20E-8	2,62E-9	1,17E-8	2,67E-9	7,26E-9	-1,10E-9
Conj. 5	9,09E-9	1,73E-8	6,61E-9	5,10E-9	9,52E-9	1,16E-9
Valor Médio	9,40E-9	7,63E-9	7,25E-9	9,11E-9	8,36E-9	

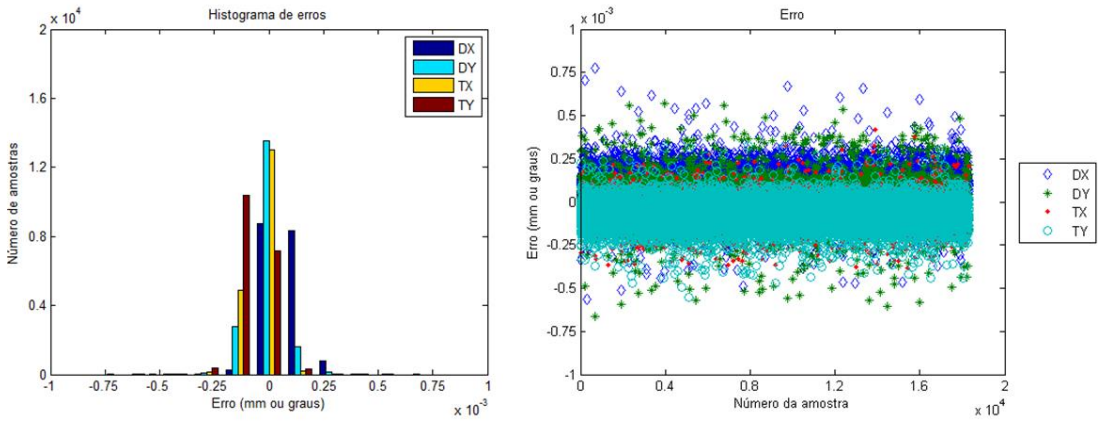

Figura 43 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia III, treinamento 1.

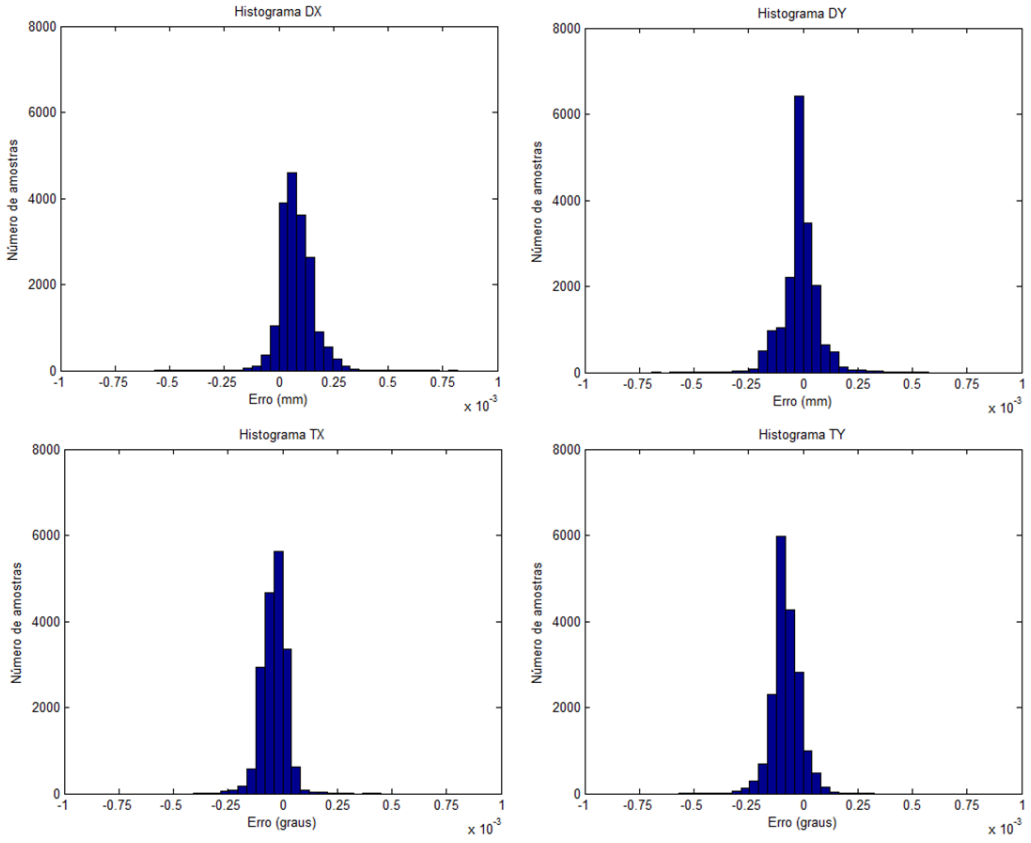

Figura 44 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia III, treinamento 1.

Estratégia IV:

Os resultados apresentados a seguir correspondem a uma rede com 32, 16 e 4 neurônios respectivamente. As camadas escondidas possuem função de ativação do tipo logística e a camada de saída possui função de ativação linear.

A Tabela 17 apresenta o EQM para cada uma das 4 saídas e o valor total diante das amostras de teste. O valor médio para cada saída e os desvios dos valores do EQM total em cada treinamento em relação ao seu valor médio também são apresentados. A Tabela 18 apresenta a porcentagem das amostras com erro, em módulo, inferior a 0,005mm ou 0,005°. A Figura 45 apresenta o EQM em função do número de épocas de treinamento para o treinamento 1. O valor mínimo atingido durante treinamento foi 1,76E-6.

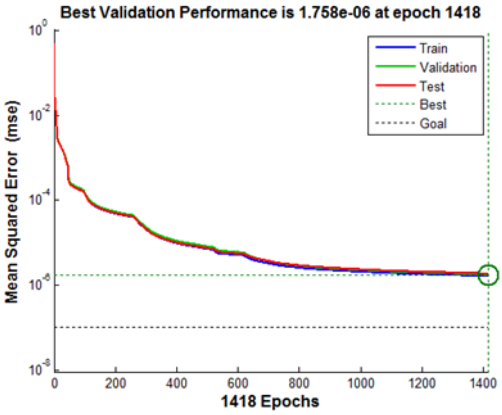

Figura 45 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada com as amostras referentes à estratégia IV, treinamento 1.

Tabela 17 – EQM para cada uma das quatro saídas da RNA e valor total, valores médios e desvio dos valores de EQM total, em cada treinamento, em relação ao seu valor médio, treinamentos com as amostras referentes à estratégia IV – amostras de teste.

ESTRATÉGIA IV						
Treinamento	**DX**	**DY**	**EQM TX**	**TY**	**TOTAL**	**Desvio (EQM TOTAL)**
Conj. 1	1,46E-5	1,37E-5	3,55E-7	8,06E-7	7,35E-6	-3,83E-6
Conj. 2	2,17E-5	3,40E-5	1,17E-6	6,04E-7	1,44E-5	3,22E-6
Conj. 3	1,60E-5	2,90E-5	1,82E-6	3,93E-7	1,18E-5	6,20E-7
Conj. 4	1,54E-5	2,01E-5	6,43E-7	4,31E-7	9,15E-6	-2,03E-6
Conj. 5	2,46E-5	2,69E-5	7,04E-7	6,31E-7	1,32E-5	2,02E-6
Valor Médio	1,85E-5	2,47E-5	9,38E-7	5,73E-7	1,12E-5	

Tabela 18 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das quatro saídas da RNA e valor total, treinamentos com as amostras referentes à estratégia IV – amostras de teste.

ESTRATÉGIA IV					
Treinamento	**DX**	**DY**	**TX**	**TY**	**TOTAL**
Conj. 1	95,3	91,7	99,8	99,8	96,6
Conj. 2	84,1	83,1	99,5	99,7	91,6
Conj. 3	91,3	83,6	99,4	99,7	93,5
Conj. 4	89,7	88,4	99,7	99,9	94,5
Conj. 5	86,1	87,2	99,6	99,8	93,2
Valor médio	89,3	86,8	99,6	99,8	93,9

A Figura 46 apresenta o histograma de erros e um gráfico de erro em função do número da amostra para as quatro saídas da RNA para o treinamento 1. A Figura 47 apresenta os histogramas de erros para cada uma das quatro saídas no caso do treinamento 1.

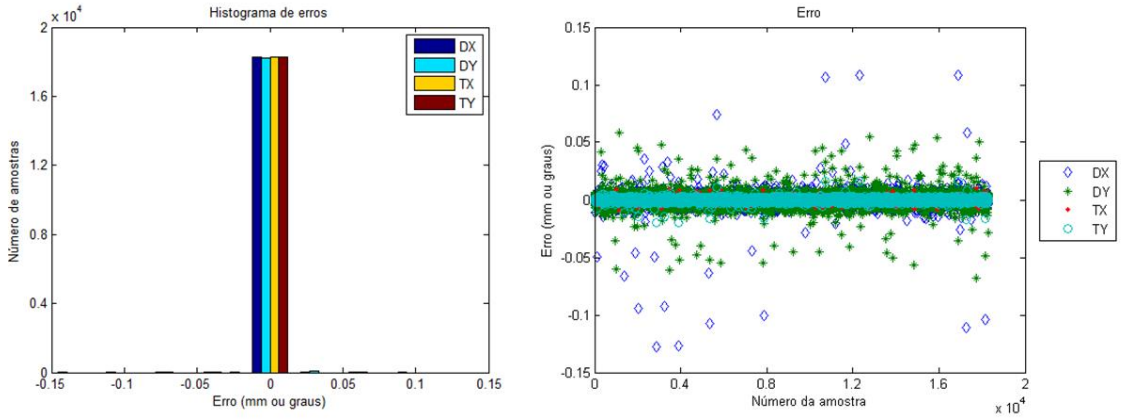

Figura 46 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras referentes à estratégia IV, treinamento 1.

Comparações entre as estratégias I e IV e as estratégias II e III possibilitam avaliar o desempenho da rede na previsão dos erros de alinhamento quando coeficientes como inclinação e *defocus* são apresentados à rede. A análise do desempenho da rede nestas duas condições pode esclarecer se estes coeficientes necessitam ou não ser apresentados a ela.

O que pôde ser observado foi que entre as estratégias II e III, o desempenho da rede, no que diz respeito ao EQM atingido, foi ligeiramente melhor na estratégia III. Contudo, o desempenho da rede em ambos os casos foi satisfatório, fato que pode ser observado no gráfico do valor do EQM em função da época de treinamento e nos histogramas de erros e gráficos de erro em função do número da amostra.

Observa-se que no caso das amostras de teste, o EQM médio ficou da ordem de 8,4E-9 no caso da estratégia III e 8,8E-9 no caso da estratégia II, e apenas 8 neurônios na primeira camada escondida e 16 na segunda camada escondida foram suficientes para a convergência do treinamento nestas duas estratégias.

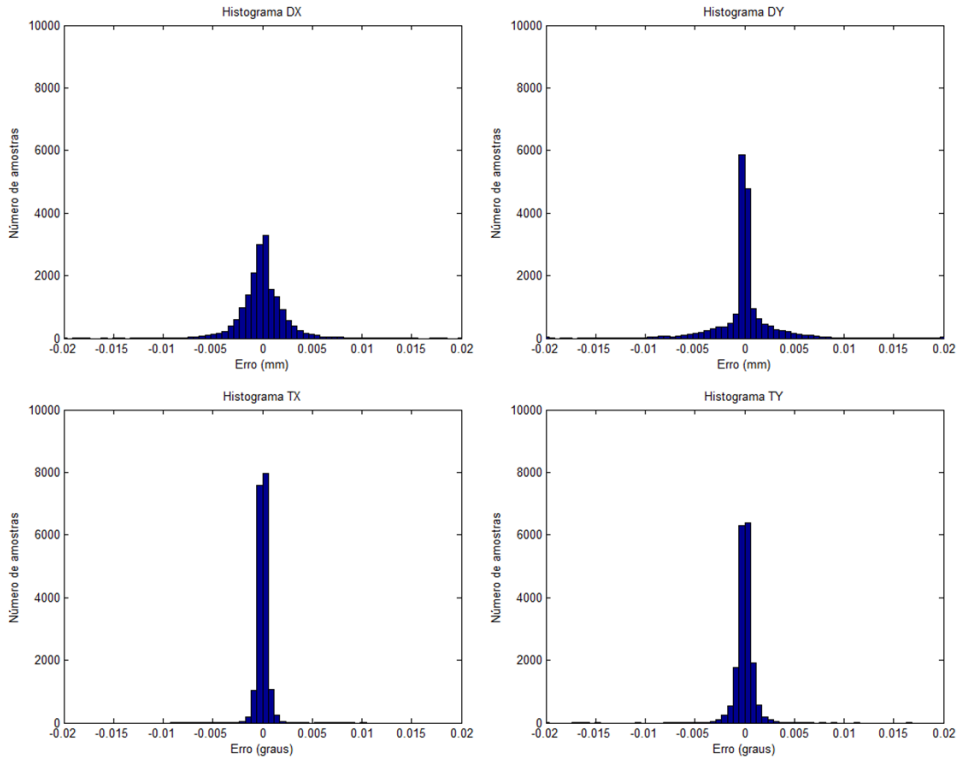

Figura 47 – Histogramas de erros para cada uma das quatro saídas para a RNA treinada com as amostras referentes à estratégia IV, treinamento 1.

Os coeficientes Z_2 e Z_3, relacionados à inclinação em *x* e *y*, e Z_4 relacionado a *defocus* não ocasionaram uma melhora significativa do desempenho. Os coeficientes Z_7 e Z_8 são relacionados a coma e à inclinação (4, 44, 45), e os coeficientes Z_5 e Z_6 são relacionados a astigmatismo e *defocus.* Os resultados indicam que dados de entrada formados por estes quatro coeficientes são suficientes para que a RNA consiga estimar o estado de alinhamento do sistema.

Para as redes treinadas com as estratégias I e IV, observa-se que em ambos os casos o desempenho é inferior ao das estratégias II e III e o EQM atingido durante treinamento foi de 2,6E-5 e 1,1E-5, respectivamente, cerca de quatro ordens de grandeza maiores que no caso das estratégias II e III em que medidas dos coeficientes em outras quatro posições no campo são utilizadas, além daquelas sobre o eixo.

A estratégia IV, em que valores de coeficientes relacionados a inclinação e *defocus* foram utilizados no treinamento da rede apresentou desempenho ligeiramente melhor, embora o ganho não seja muito significativo.

Observa-se que nas estratégias I e IV a rede apresenta melhor desempenho quando desalinhamentos do tipo inclinação são estimados, apresentando maior erro para as saídas que dizem respeito a descentralizações.

Quando os coeficientes Z_2, Z_3 e Z_4 são adicionados aos dados de entrada, ocorre um considerável aumento no número de dados de entrada, principalmente no caso da estratégia II totalizando em 35 dados ao invés de 20, no caso da estratégia III.

Comparando-se redes treinadas com as estratégias I e III e com as estratégias II e IV, pode-se concluir que quando os valores dos coeficientes medidos em mais de um ponto no campo são apresentados à rede, o erro na estimativa dos desalinhamentos é menor. Ganhos de até quatro ordens de grandeza no valor do EQM para as amostras de teste foram observados quando os valores dos coeficientes em cinco pontos no campo são considerados.

Quando a medida da frente de onda em pontos fora do eixo se torna muito complicada, treinamentos da rede neural, usando-se apenas coeficientes medidos no eixo óptico, podem ser realizados. Entretanto, as estimativas dos desalinhamentos se tornam muito mais precisas quando medidas fora do eixo são consideradas.

Conclui-se que no caso do sistema analisado, a estratégia que apresentou desempenho satisfatório aliado a um menor número de dados de entrada, quando comparado à estratégia II, que apresentou desempenho similar, foi a estratégia III, constituída pelos coeficientes Z_5, Z_6, Z_7 e Z_8 relativos a astigmatismo e coma, e consiste de medidas no eixo e em outros quatro pontos no campo.

Como os efeitos dos erros de alinhamento são diferentes de um sistema para outro, análises similares devem ser realizadas especificamente para cada sistema que se deseja alinhar, visando avaliar qual o melhor conjunto de dados de entrada, de modo que este contenha informação suficiente para que o treinamento da RNA apresente desempenho satisfatório.

5.3 Efeitos de erros de forma nas superfícies dos elementos

5.3.1 Modelagem óptica das superfícies com erros de forma

Dois parâmetros são importantes quando se trata de erros de superfície:

- magnitude do erro: que pode ser dado, por exemplo, pelo valor do erro RMS introduzido na superfície;
- frequência espacial das irregularidades.

O Zemax possui ferramentas (TEZI, TEXI), usadas nas análises de tolerância, capazes de gerar padrões de irregularidades na superfície de um determinado elemento do sistema que simulam os erros de superfície. A função TEZI funciona da seguinte forma: o software substitui a superfície original do sistema por uma do tipo *Zernike Standard Sag*, que possui o mesmo formato esférico padrão e adicionalmente, os coeficientes de Zernike representam o desvio dessa superfície em relação ao formato nominal. Esta ferramenta assume as seguintes características (59):

- a superfície original do elemento é uma superfície do tipo *Standard*, usada frequentemente para modelar uma superfície esférica, ou *Even aspheric*, usada para modelar superfícies asféricas;
- os polinômios de Zernike podem representar, de forma satisfatória, os erros de forma. Segundo diversos autores, esta é uma boa suposição nos casos em que interferômetros são usados para testar essas superfícies, uma vez que os softwares dos interferômetros são capazes de representar os erros entre o *sag* original da superfície e o produzido por um determinado número de termos.

Por meio dessa ferramenta, definindo-se qual a amplitude do erro RMS das irregularidades da superfície e qual o menor e o maior coeficiente de Zernike a ser considerado, o programa gera, de forma aleatória, diferentes padrões de irregularidades que resultam no valor RMS pré-estabelecido. Essas superfícies

irregulares geradas podem ser salvas e usadas, posteriormente, para a extração de dados de entrada para a RNA de sistemas contendo superfícies com erros aleatórios de fabricação.

Quando termos de baixas ordens são usados, as irregularidades apresentarão baixas frequências. Se termos de altas ordens são usados, haverá irregularidades de frequências altas.

A escolha da faixa de termos a ser usada pode variar de um fabricante para outro. Uma maneira de definir o intervalo de termos ideal a ser considerado, de acordo com sugestão de Kim (59), é avaliar peças produzidas por um determinado fabricante, por meio de medidas interferométricas, e ajustar os polinômios de Zernike à frente de onda medida, de forma a definir o menor e o maior termo para a obtenção de um bom ajuste.

Utilizando-se a ferramenta TEZI, foi possível gerar diferentes deformações superficiais nos espelhos primário e secundário do telescópio *Cassegrain* avaliado.

Para a definição do intervalo de polinômios de Zernike (*Standard*) a ser utilizado, optou-se por escolher dois intervalos de 5 a 22 e de 5 a 231. A partir da análise das superfícies de diversas lentes fabricadas na Opto eletrônica S.A. constatou-se que nelas, o primeiro intervalo de valores adotado é o mais adequado.

Os termos 2 e 3 estão relacionados à inclinação. O termo 4 corresponde a *defocus*. Estes três termos não foram considerados, a principio, nas análises de erros de forma. A Figura 48 apresenta a frente de onda de uma mesma superfície sem (a) e com (b) os termos de inclinação. Observa-se que sua influência é grande nos valores RMS e P-V medidos da superfície.

De acordo com a descrição dos termos, dada por Wyant (50) e com a ordenação dos polinômios de Zernike *STANDARD* do software Zemax (45) , os termos 5 e 6 referem-se ao astigmatismo primário, 7 e 8 ao coma primário, 9 e 10 referem-se ao *trefoil* e o 11 à aberração esférica primária. Os termos 12 e 13 referem-se ao astigmatismo secundário, 14 e 15 referem-se a *tetrafoil*, 16 e 17 referem-se ao coma secundário, 18 e 19 a *trefoil* secundário, 20 e 21 referem-se a *pentafoil* e, por fim, o termo 22 refere-se à aberração esférica secundária.

Dessa forma, no primeiro intervalo escolhido, erros comáticos, astigmáticos, esféricos e os devido às demais aberrações citadas serão combinados de forma

aleatória, causando deformações com valor máximo de erro RMS nas superfícies dos elementos ópticos estipulado. O segundo intervalo abrange termos de aberrações de ordens mais altas.

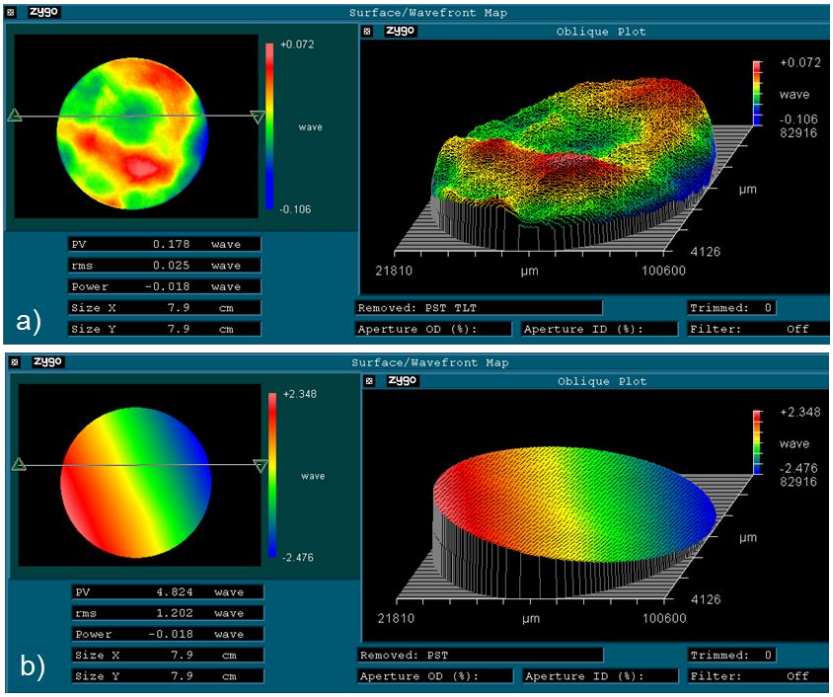

Figura 48 – a) Frente de onda da superfície de uma lente em que *piston* e inclinação (*tilt*) foram descontados da imagem; b) Frente de onda da mesma superfície em que apenas *piston* foi descontado. Fonte: Opto Eletrônica S.A.

O valor do erro RMS a ser considerado foi escolhido baseado na análise da frente de onda refletida das superfícies dos elementos ópticos fabricados na Opto Eletrônica S.A. Os elementos analisados apresentavam tolerâncias de fabricação de $0{,}25\ \lambda$ de P-V (*peak-to-valey*) e apresentavam superfícies esféricas, sendo fabricados por meio dos métodos convencionais (Figura 49). O valor de erro RMS encontrado nestas superfícies, para um comprimento de onda de 632,8 nm, foi de $0{,}04\ \lambda$ a $0{,}045\ \lambda$, quando o valor de P-V era da ordem do valor especificado, o que corresponde a cerca de 0,03 µm.

Devido ao método de medida, o valor do erro a ser considerado na superfície é de 0,015 µm (ou 1,5E-5 mm) já que o erro da frente de onda tem seu valor dobrado devido a reflexão na superfície (75) e as figuras apresentadas na Figura 49 foram obtidas com o fator de escala do interferômetro ajustado para 1.

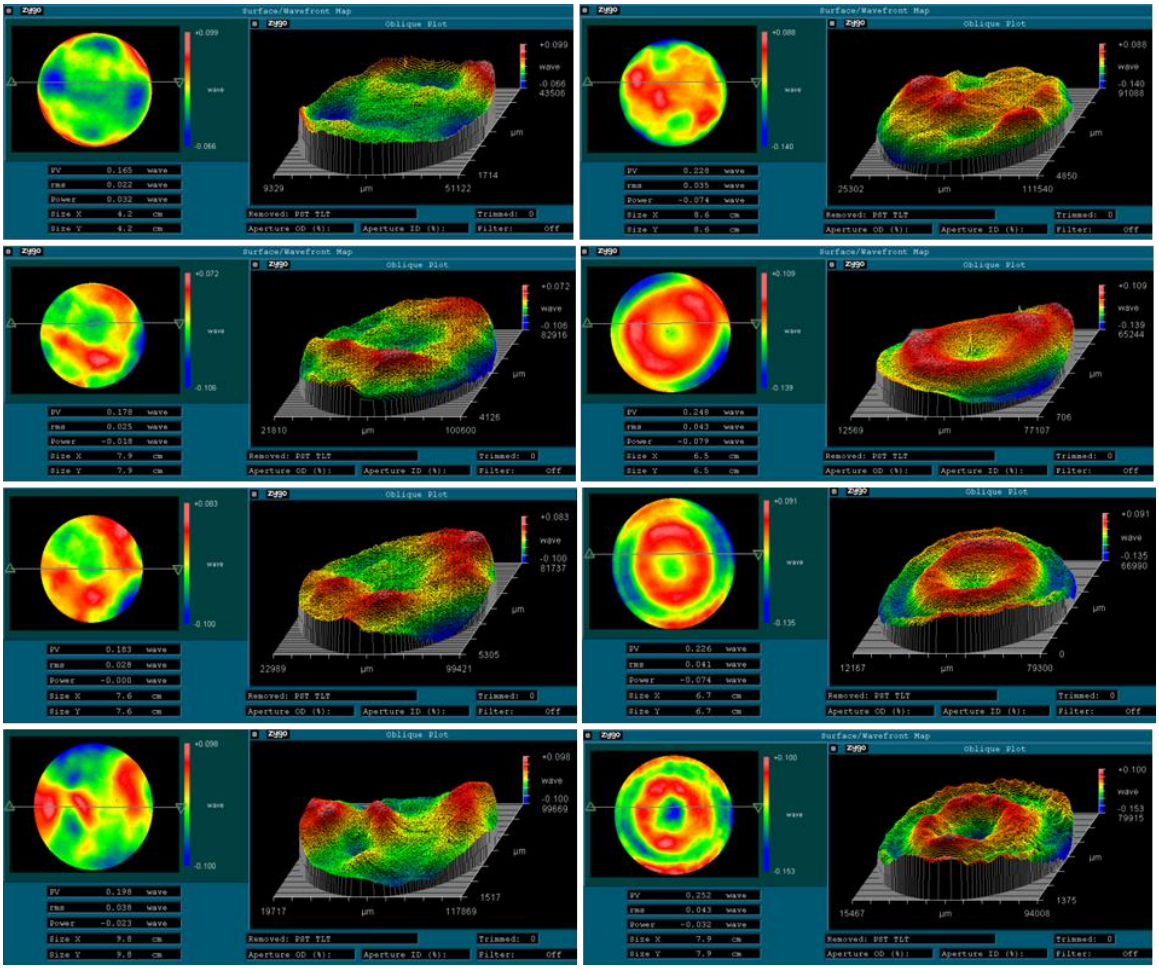

Figura 49 – Frentes de onda de superfícies de diferentes lentes cujas especificações de fabricação para erros superficiais eram de 0,25 λ de P-V. Fonte: Opto Eletrônica S.A.

Considerando-se os coeficientes e o valor de erro RMS descritos anteriormente, combinações aleatórias de erros superficiais foram criadas para cada elemento do sistema.

A Figura 50 apresenta exemplos dos padrões das deformações geradas nas superfícies de ambos os espelhos, quando os coeficientes de 5 a 22 foram considerados; a Figura 51 apresenta exemplos de padrões de deformação gerados nas superfícies dos espelhos quando os coeficientes de 5 a 231 foram considerados. Apenas as deformações são apresentadas nestas figuras, o *sag* da superfície original foi descontado para melhor visualização das irregularidades.

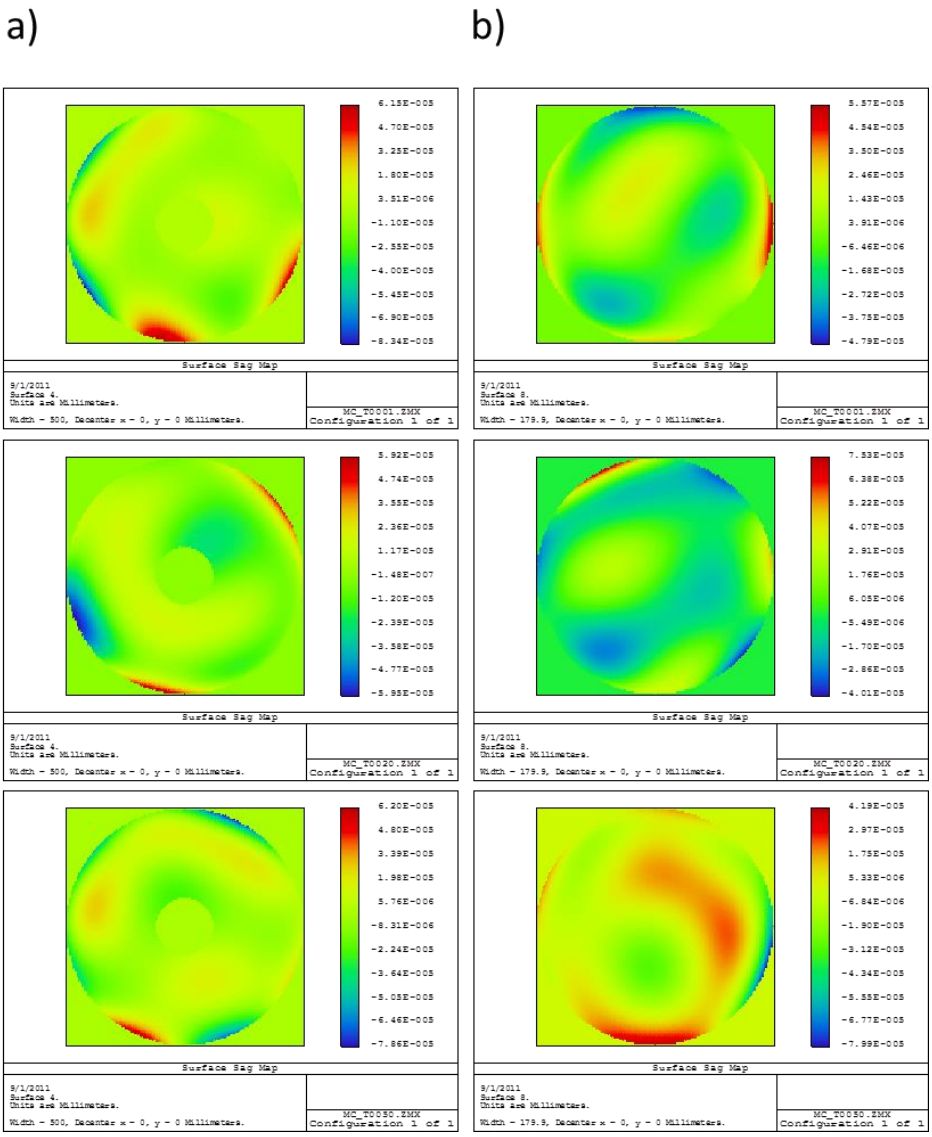

Figura 50 – Exemplos de erros de forma gerados nas superfícies dos espelhos (a) primário e (b) secundário considerando coeficientes de Zernike (*Standard*) de 5 a 22 e RMS=1,5E-5 mm.

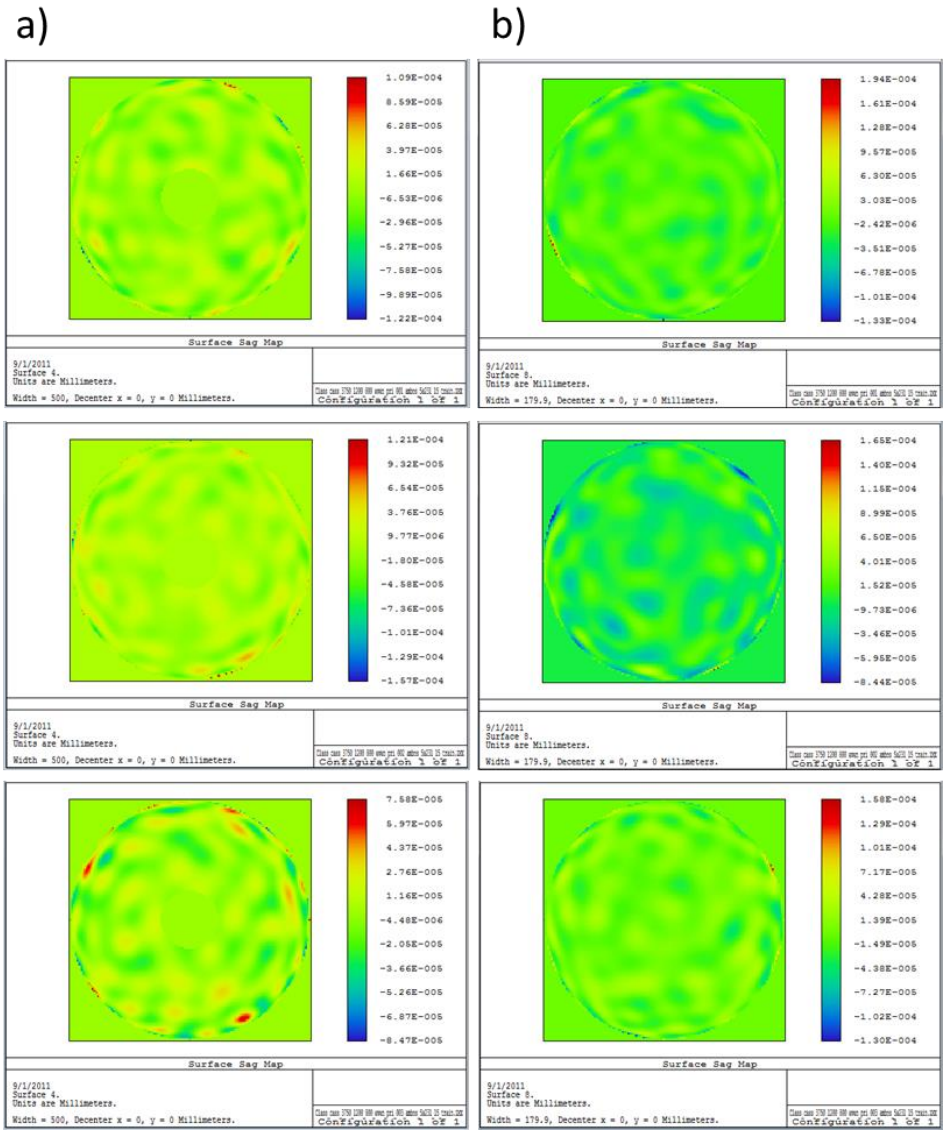

Figura 51 – Exemplos de erros de forma gerados nas superfícies dos espelhos (a) primário e (b) secundário considerando coeficientes de Zernike (*Standard*) de 5 a 231 e RMS=1,5E-5 mm.

Os detalhes da análise realizada são apresentados a seguir.

5.3.2 Resultados e discussões

5.3.2.1 Avaliação do desempenho da rede treinada com dados de entrada de acordo com a estratégia III considerando apenas amostras provenientes do sistema original durante treinamento

Para avaliar o efeito que esses erros de forma causam no desempenho da rede, foi utilizada uma rede treinada com dados de entrada compostos pelos coeficientes dos polinômios de Zernike especificados na estratégia III. Durante o treinamento, apenas dados provenientes do sistema original, sem erros de forma, foram apresentados à rede.

Tabela 19 – Descrição dos intervalos de desalinhamento adotados na análise dos efeitos de deformações superficiais no desempenho da rede.

Desalinhamento	**Intervalo**
Descentralização em *x*	-1,0 a 1,0 mm
Descentralização em *y*	-1,0 a 1,0 mm
Inclinação em *x*	-1,0 a 1,0°
Inclinação em *y*	-1,0 a 1,0°
Descentralização em *x* e *y*	-1,0 a 1,0 mm
Inclinação em *x* e *y*	-1,0 a 1,0°
Descentralização em *x* e Inclinação em *x*	-1,0 a 1,0 mm -1,0 a 1,0°

Os resultados apresentados a seguir referem-se a uma rede com duas camadas escondidas, com 8 e 4 neurônios, respectivamente, e função de ativação logística. A camada de saída possui 4 neurônios com função de ativação linear.

Depois de treinada, amostras provenientes de cinco sistemas contendo superfícies com deformações com erros RMS de 1,5E-5 mm foram apresentadas à rede. Essas deformações foram geradas considerando-se dois intervalos de coeficientes dos polinômios de Zernike *Standard* de 5 a 22 e de 5 a 231. A extração dos dados foi realizada para cada sistema individualmente, tomando um tempo considerável. Por

esse motivo, no primeiro intervalo, que será avaliado mais profundamente, foram considerados os erros de alinhamento descritos na Tabela 19 e apenas descentralização e inclinação em *x* e em *y* (nos mesmos intervalos e amostragens) foram consideradas no caso do segundo intervalo de coeficientes de Zernike, analisado aqui brevemente.

Amostras extraídas destes sistemas foram tratadas da seguinte forma:

- Obtenção dos dados diferenciais: os valores dos coeficientes Z_5, Z_6, Z_7 e Z_8 para o sistema em seu estado alinhado, proveniente do sistema sem imperfeições, que é o sistema que efetivamente foi usado para treinamento da rede, foi descontado das amostras extraídas destes sistemas com deformações.

- Normalização: os coeficientes foram normalizados também considerando-se os valores mínimo e máximo de cada coeficiente provenientes do sistema original, sem imperfeições.

A Tabela 20 apresenta os valores do EQM para cada uma das quatro saídas da rede no caso de sistemas sem deformações e com deformações contendo coeficientes nos intervalos de 5 a 22 e de 5 a 231. A Figura 52 apresenta o histograma de erros e um gráfico de erro em função do número da amostra para o intervalo de 5 a 22. A Figura 53 apresenta o histograma de erros e um gráfico de erro em função do número da amostra para o intervalo de 5 a 231. As amostras de teste apresentadas à rede previamente treinada foram mantidas separadas por sistema e em ordem crescente de erro de alinhamento.

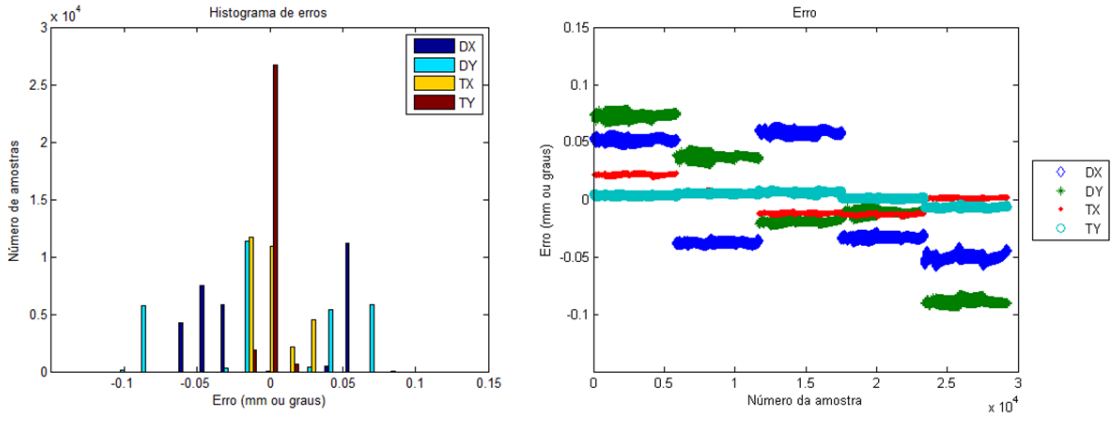

Figura 52 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras do sistema perfeito referentes à estratégia III, para amostras vindas de cinco sistemas imperfeitos com deformações contendo coeficientes de Zernike de 5 a 22.

Tabela 20 – EQM para cada uma das quatro saídas da RNA diante de amostras de teste do sistema sem deformações e de sistemas com imperfeições geradas com coeficientes de Zernike nos intervalos de 5 a 22 e de 5 a 231.

	DX	DY	TX	TY
EQM (sistema sem deformações)	3,8E-7	3,3E-7	4,7E-8	6,3E-8
EQM (coef. Zernike de 5 a 22)	2,3E-3	3,1E-3	1,7E-4	2,6E-5
EQM (coef. Zernike de 5 a 231)	9,3E-4	3,7E-4	5,1E-5	9,4E-5

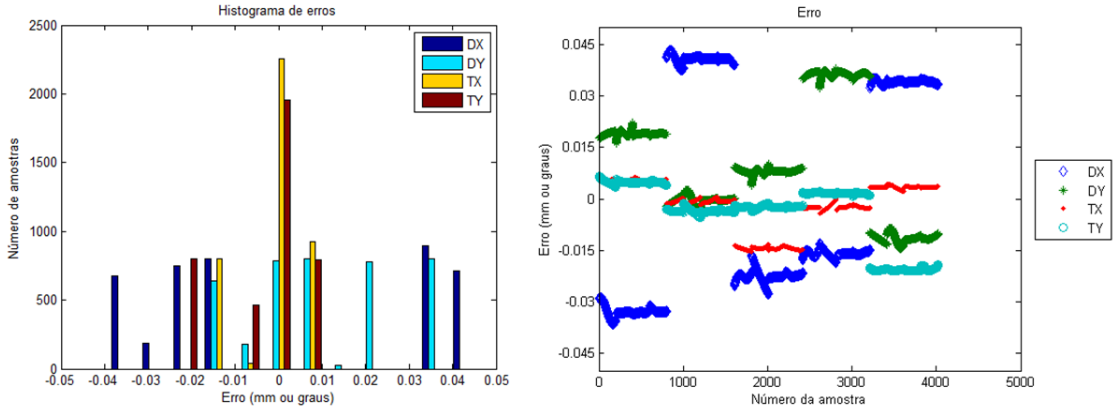

Figura 53 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede treinada com as amostras do sistema perfeito referentes à estratégia III, para amostras vindas de sistemas imperfeitos com deformações contendo coeficientes de Zernike de 5 a 231.

Observa-se nas tabelas apresentadas anteriormente que o desempenho da rede é degradado quando amostras provenientes de sistemas contendo deformações superficiais são apresentadas a ela em ambos os intervalos de coeficientes considerados.

Para entender o motivo, algumas saídas da rede foram analisadas (Figura 54 e Figura 55), além disso, os coeficientes Z_5, Z_6, Z_7 e Z_8 foram analisados comparando-se os valores extraídos do sistema original com os extraídos de dois sistemas contendo deformações superficiais (Figura 56 a Figura 59).

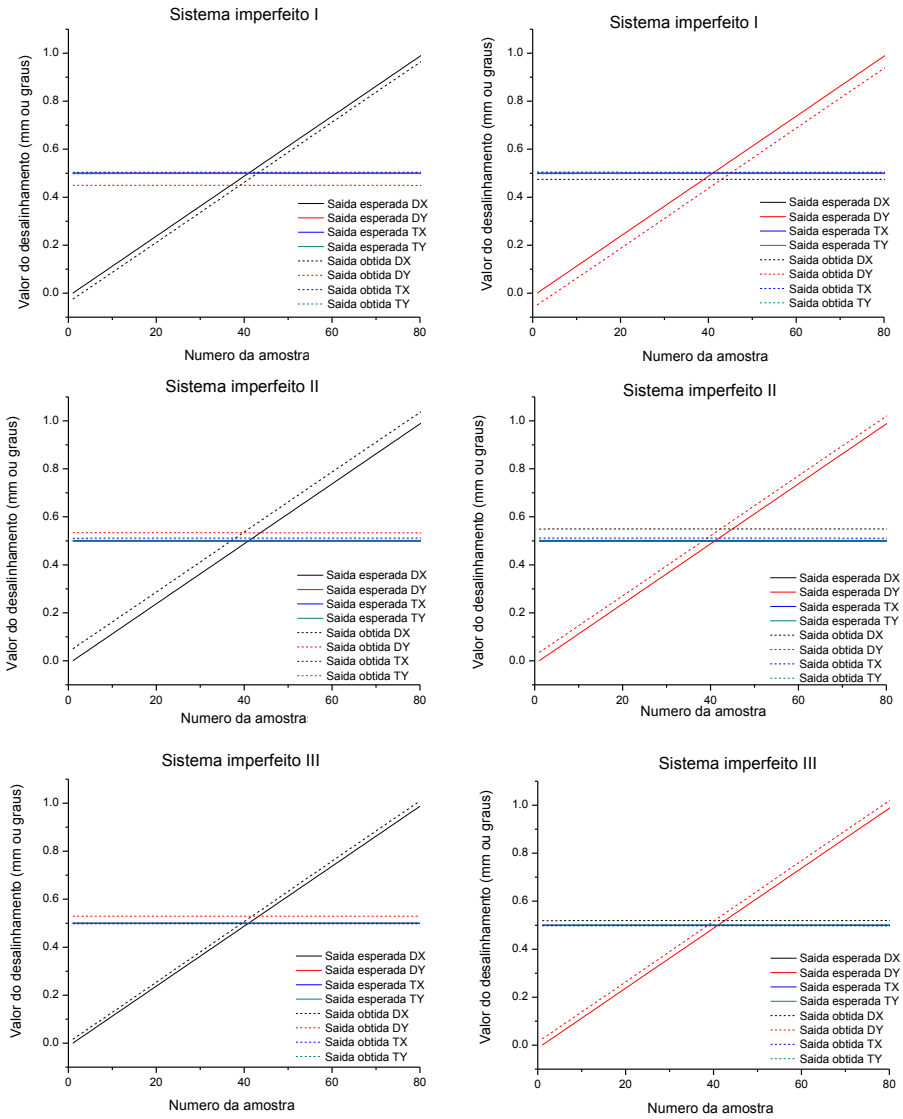

Figura 54 – Saídas da rede treinada apenas com dados provenientes do sistema perfeito diante da apresentação das amostras extraídas de sistemas imperfeitos. Casos de descentralização em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada.

Observa-se nos gráficos da Figura 54 e da Figura 55, que mostram a saída esperada e a obtida nos casos de três sistemas imperfeitos, com deformações aleatórias, que ocorre um deslocamento nos valores obtidos como saída da rede quando comparados aos valores esperados. Este deslocamento das curvas de saída esperada e saída obtida é diferente em cada caso de deformação superficial e é, de modo geral, maior no caso de descentralização. Apenas desalinhamentos em um dos graus de liberdade são analisados em cada gráfico das figuras citadas.

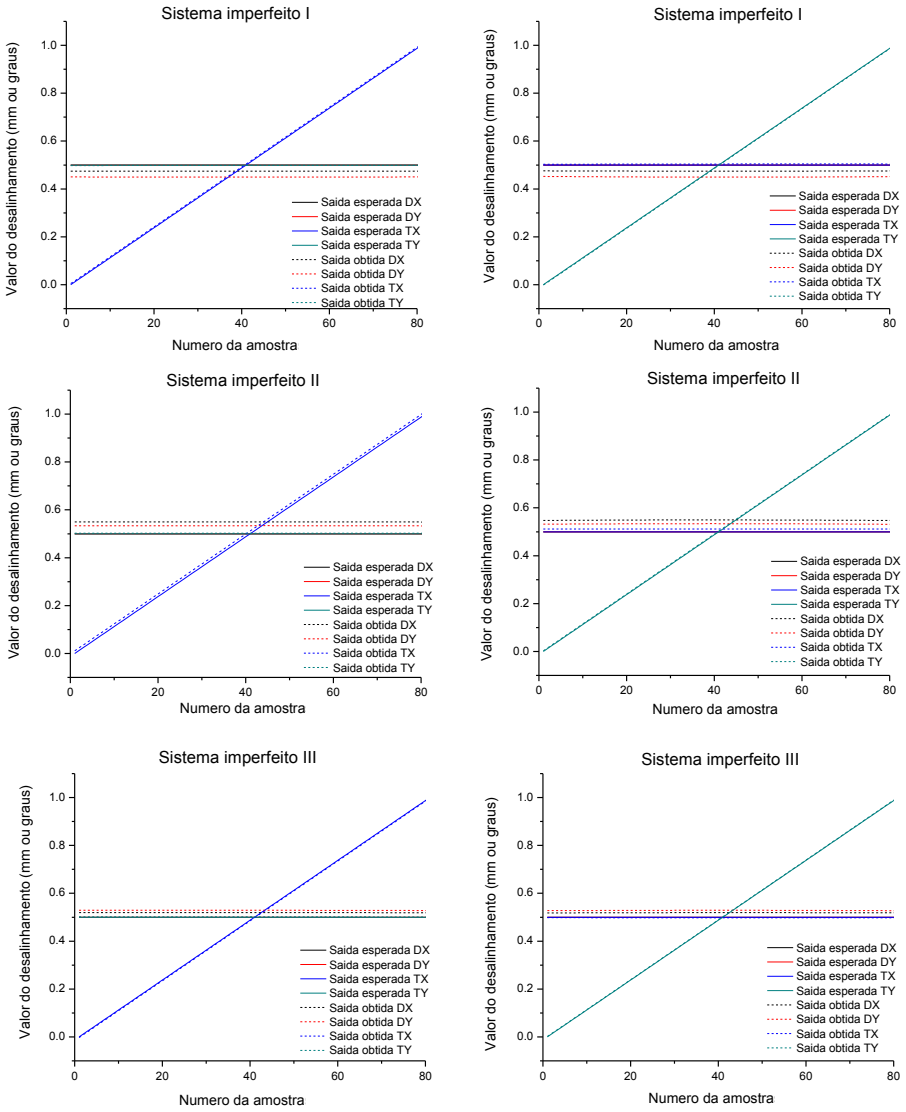

Figura 55 – Saídas da rede treinada apenas com dados provenientes do sistema perfeito diante da apresentação das amostras extraídas de sistemas imperfeitos. Casos de erros de alinhamento do tipo inclinação em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada.

Este efeito também pode ser observado nos gráficos de erro em função do número da amostra (Figura 52 e Figura 53) onde as amostras de cada sistema contendo imperfeições superficiais resultaram em erros da mesma ordem de grandeza formando focos de desvios distintos. Mais uma vez é possível observar que esses erros são maiores no caso de descentralização e que são diferentes para cada uma das quatro saídas.

A Figura 56 apresenta as curvas de Z_5 em função do desalinhamento para o sistema original (perfeito) e dois dos sistemas contendo deformações superficiais nos espelhos com RMS da ordem de 1,5E-5 mm em cada superfície.

Os coeficientes foram extraídos em cinco pontos no campo, conforme pode ser observado na legenda.

Nesse gráfico, os dados estão normalizados no intervalo entre 0 e 1. Apenas desvios de alinhamento puros, do tipo DX, DY, TX ou TY, foram considerados.

Observa-se que ocorre um deslocamento das curvas para os sistemas com deformações superficiais em comparação com as curvas de Z_5 extraídas do sistema original. O comportamento de Z_5 em função do desvio de alinhamento não é afetado pelas deformações. Observa-se também que para os sistemas com imperfeições, no caso de desalinhamentos do tipo inclinação, para os quais este coeficiente apresenta maior variação com o desalinhamento, alguns desvios de alinhamento resultam em coeficientes com valor maior que 1, ou seja, fora da faixa de normalização (0 a 1). Gráficos similares são apresentados na Figura 57 para o coeficiente Z_6, na Figura 58 para Z_7 e na Figura 59 para Z_8.

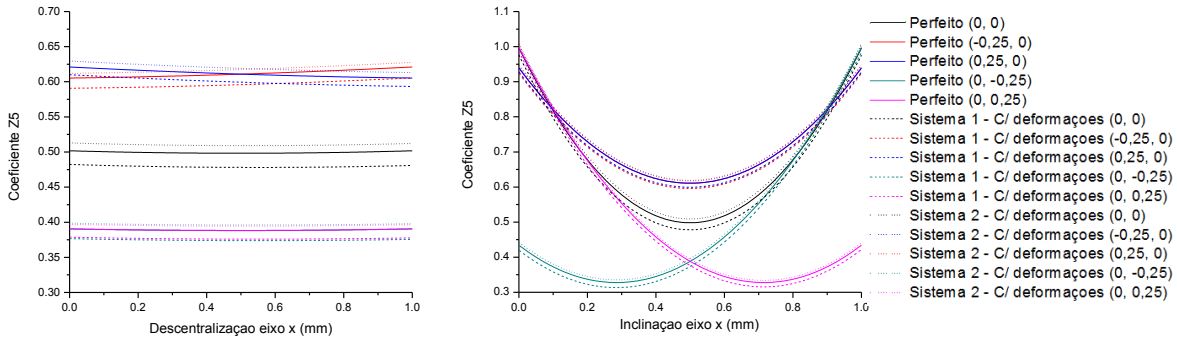

Figura 56 – Coeficiente Z_5 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo, para o sistema original, e dois sistemas com deformações superficiais tratados conforme descrito anteriormente.

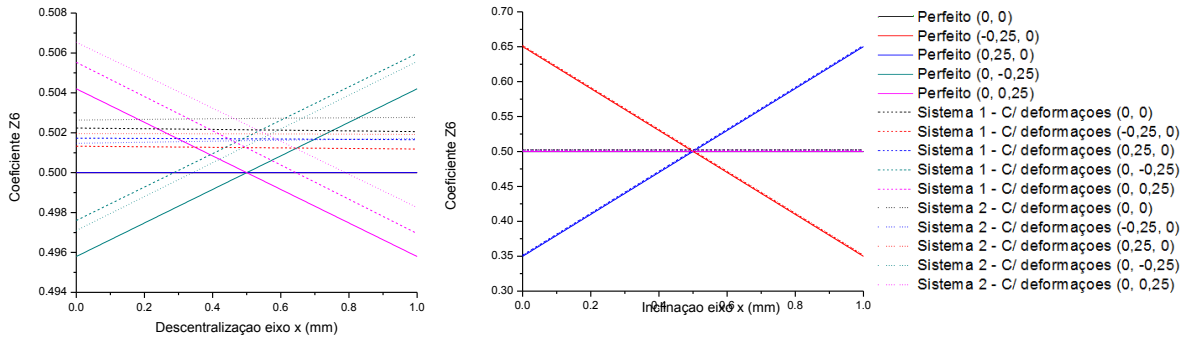

Figura 57 – Coeficiente Z_6 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo, para o sistema original, e dois sistemas com deformações superficiais.

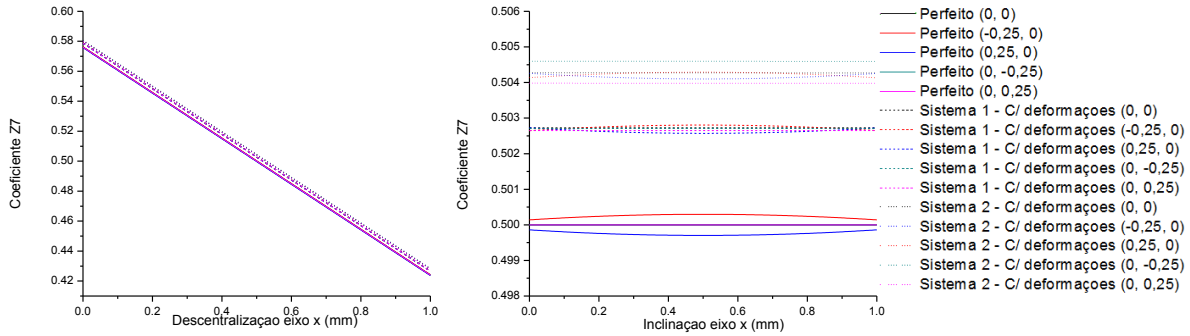

Figura 58 – Coeficiente Z_7 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campo para o sistema original, e dois sistemas com deformações superficiais.

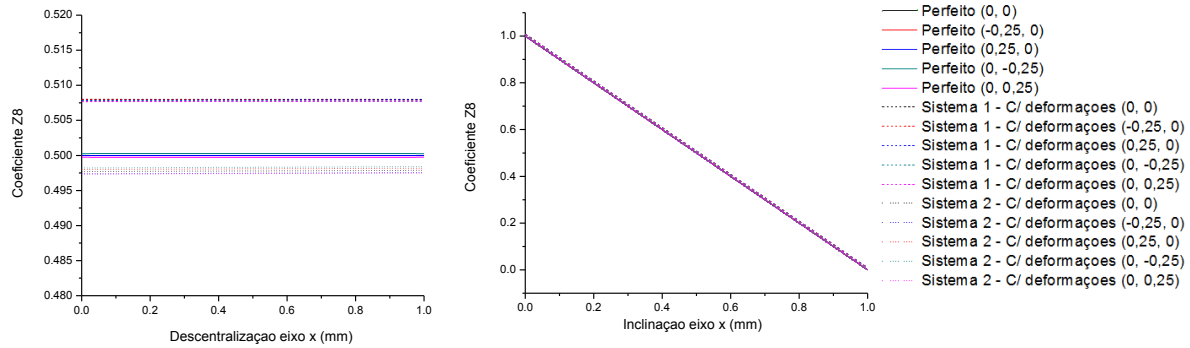

Figura 59 – Coeficiente Z_8 em função de a) descentralização e b) inclinação do espelho secundário para cinco ponto no campos para o sistema original, e dois sistemas com deformações superficiais.

5.3.2.2 Melhoramento dos valores estimados - RNA corretora

Com base nos resultados obtidos nas análises anteriores, buscou-se uma alternativa para a obtenção de estimativas mais precisas de desvios de alinhamento quando defeitos superficiais estão presentes nos elementos do sistema.

O treinamento de uma RNA considerando amostras provenientes de sistemas com erros de forma nas superfícies de seus elementos diretamente, não resultou em melhoras significativas nos erros dos valores estimados.

A Figura 60 apresenta as saídas esperada e obtida de uma rede treinada com amostras provenientes de sistemas com deformações superficiais para dados de teste provenientes dos mesmos sistemas, porém que não haviam sido apresentados à rede durante treinamento, para o caso de descentralização. Observa-se que neste caso as saídas esperadas e obtidas apresentam erros de magnitudes diferentes, dependendo da magnitude dos desalinhamentos.

Uma alternativa encontrada foi a utilização de uma segunda RNA corretora dos dados estimados a partir da RNA estimadora, treinada apenas com dados do sistema perfeito. A Figura 61 apresenta um esquemático dos dados de entrada e saída da RNA estimadora e da RNA corretora. Nessa análise, apenas dados provenientes de sistemas com deformações superficiais geradas considerando coeficientes de números de 5 a 22 foram considerados.

Após o treinamento da RNA estimadora (RNA_E), onde apenas amostras provenientes do sistema perfeito, sem deformações, foram apresentadas durante o seu treinamento, um conjunto de amostras provenientes de sistemas com deformações superficiais aleatórias com RMS de 1,5E-5 mm foi apresentado à RNA_E, para a estimativa dos erros de alinhamento presentes nesses sistemas (Figura 62).

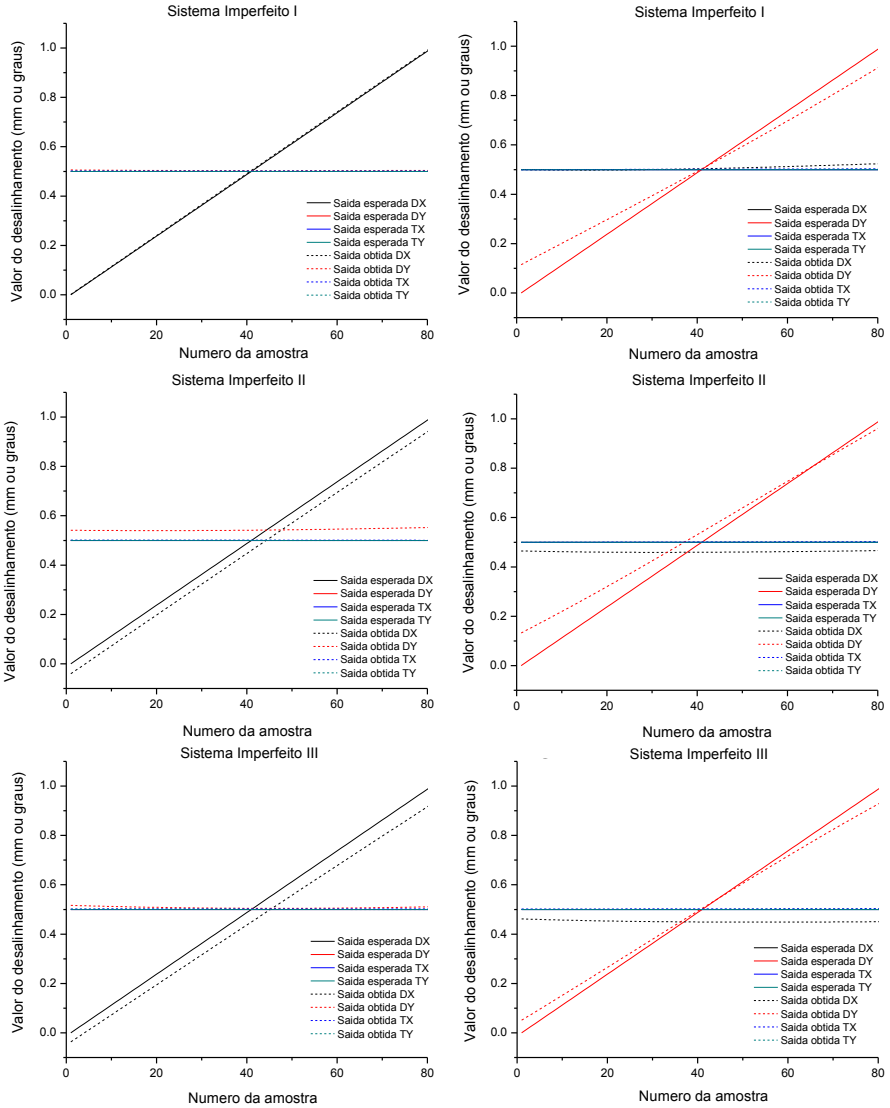

Figura 60 – Saídas da rede treinada com dados provenientes de sistemas com deformações superficiais diante da apresentação das amostras extraídas destes mesmos sistemas imperfeitos que não haviam sido apresentados durante treinamento. Casos descentralização em *x* (gráficos à esquerda) e em *y* (gráficos à direita). As saídas de três sistemas com diferentes imperfeições, aleatórias, são apresentadas e comparadas com a saída esperada.

A RNA_E considerada nesta análise possui 8 e 16 neurônios, respectivamente, nas camadas escondidas e função de ativação tangente hiperbólica. A camada de saída tem função de ativação linear.

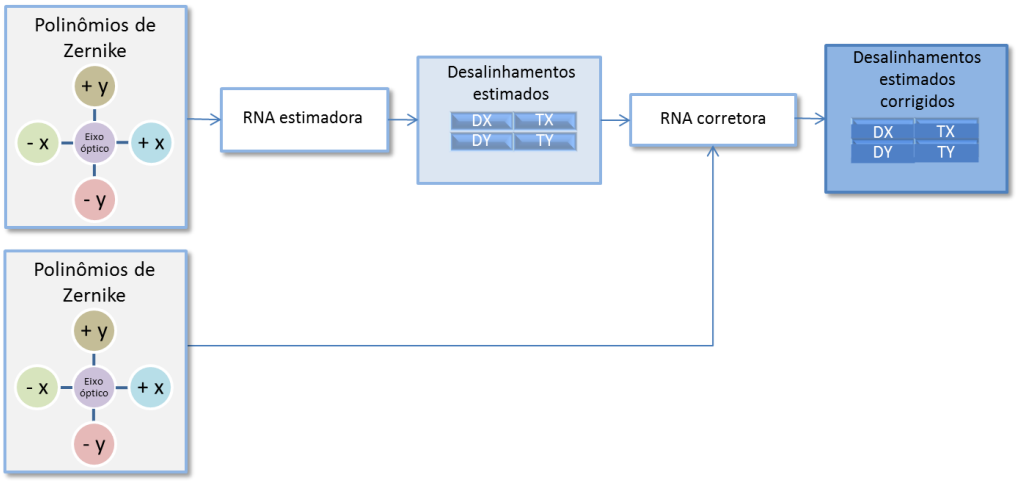

Figura 61 – Esquema ilustrativo das entradas e saídas da RNA estimadora e da RNA corretora.

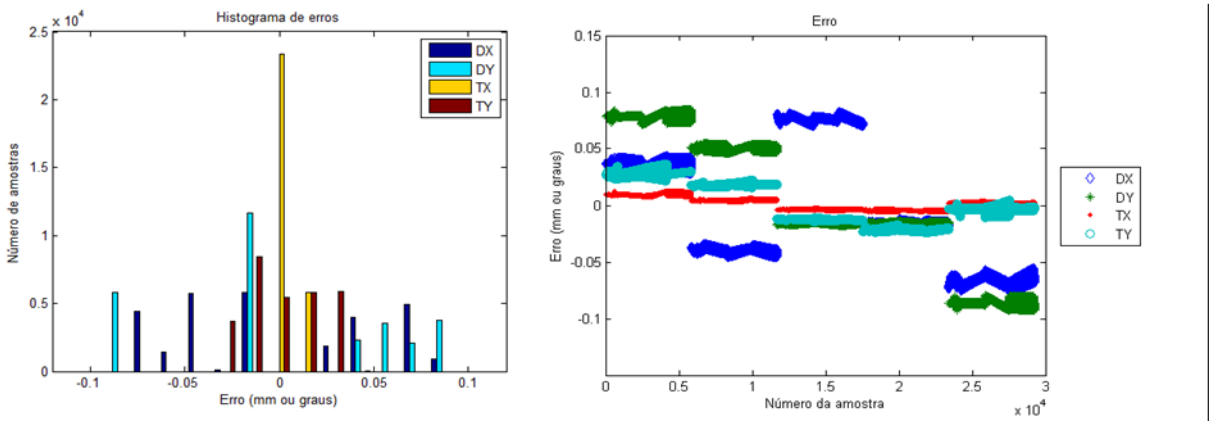

Figura 62 – Histograma de erros e gráfico de erro em função do número da amostra para as saídas da RNA_E, treinada com dados extraídos de um sistema perfeito, diante de amostras de cinco sistemas com deformações superficiais.

Os valores estimados, que constituem as saídas da RNA_E, juntamente com os coeficientes dos polinômios de Zernike (não necessariamente os mesmos coeficientes apresentados à RNA_E, ou extraídos nas mesmas posições no campo) formam agora o conjunto de dados de entrada da RNA corretora (RNA_C). Diversas combinações foram testadas. Algumas delas resultaram em alguma melhora nos dados estimados, outras não contribuiram para a melhora dos resultados.

Os dados de saída esperada dessa segunda rede neural são os mesmos da primeira, uma vez que se busca melhorar a precisão dos dados estimados. Foram considerados 45 sistemas contendo deformações aleatórias com RMS de 1,5E-5 mm geradas considerando coeficientes de Zernike entre 5 e 22. Dentre as amostras

extraídas dos sistemas, 70% foram usadas no treinamento da RNA_C, 15% foram usadas na validação e os outros 15% para teste de desempenho desta rede.

A Figura 63 apresenta o gráfico de EQM em função da época de treinamento da RNA_C. A Figura 64 apresenta o histograma de erros e um gráfico de erro em função do número da amostra para novas amostras referentes aos 15% dos dados extraídos dos mesmos 45 sistemas considerados durante o treinamento. Nas figuras apresentadas a seguir, os dados de entrada da RNA_C são compostos pelas saídas da RNA_E e pelo coeficiente Z_9 extraído em quatro posições no campo. Os resultados referem-se a RNA_c com 16 e 24 neurônios, respectivamente, e função de ativação tangente hiperbólica nas camadas escondidas. A camada de saída possuia função de ativação linear.

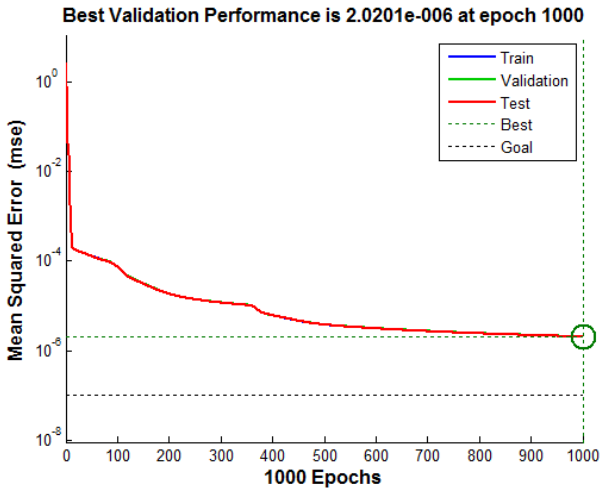

Figura 63 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede corretora treinada.

Para que fosse possível avaliar a melhoria na estimativa dos erros de alinhamento, o mesmo conjunto de amostras de teste, contendo amostras de 5 sistemas que não fazem parte do conjunto de treinamento, com desalinhamentos ordenados, cujos resultados da RNA_E foram apresentados na Figura 62, foi agora apresentado a essa rede e os erros referentes aos novos valores de desalinhamento estimados são apresentados no histograma de erros e no gráfico de erro em função do número da amostra na Figura 65. A Tabela 21 apresenta o EQM para cada saída da RNA_C para o caso de amostras referentes a sistemas com deformações superficiais já conhecidos pela RNA. O EQM total das amostras de teste foi de 1,1E-06.

Tabela 21 – EQM para cada uma das quatro saídas da RNA quando são apresentadas amostras provenientes de sistema com elementos contendo erros de superfície já conhecidos e desconhecidos pela rede.

	DX	DY	TX	TY
EQM (novas amostras de sistemas já apresentados à rede durante treinamento)	1,3E-06	1,4E-06	6,5E-07	1,1E-06
EQM (amostras de sistemas diferentes dos considerados no treinamento)	2,4E-3	1,9E-2	9,3E-5	9,2E-4

A Tabela 21 apresenta também os valores do EQM para cada saída da RNA_C para o caso de amostras referentes a sistemas com deformações superficiais desconhecidos pela RNA. O EQM total das amostras de teste foi de 5,6E-3.

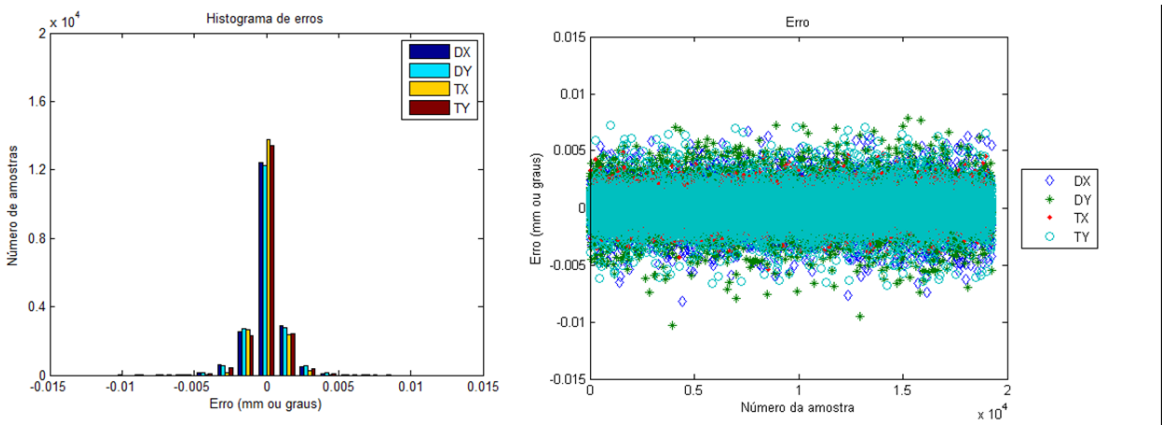

Figura 64 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede após passar pela rede corretora, para amostras vindas dos 45 sistemas imperfeitos cujas amostras foram apresentadas à RNA durante treinamento, porém estes são dados de teste que correspondem aos 15% dos dados que não foram apresentados durante treinamento da RNA corretora.

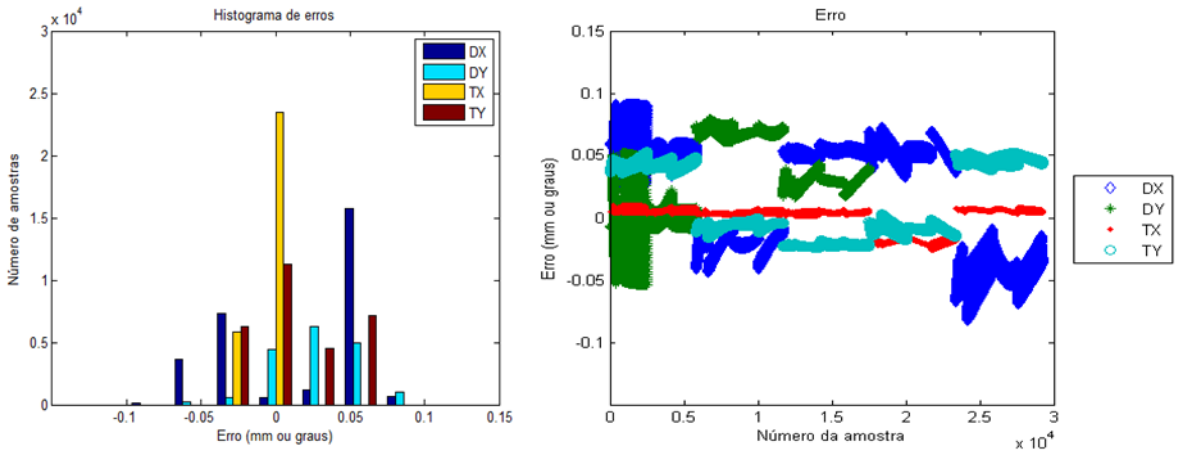

Figura 65 – Histograma de erros e gráfico de erro em função do número da amostra para as quatro saídas da rede após passar pela rede corretora, para amostras vindas dos 5 sistemas imperfeitos e que não foram considerados durante treinamento da RNA corretora.

O uso de uma segunda rede corretora (RNA_C) dos dados estimados pela RNA_E resultou em valores finais estimados com erros, em módulo, menores que 0,015 mm ou graus para desalinhamentos dos tipos descentralização ou inclinação, respectivamente, no caso da rede apresentada, para amostras inéditas provenientes de sistemas dos quais as amostras de treinamento foram extraídas.

Os erros são menores que 0,1 mm ou graus para desalinhamentos do tipo descentralização ou inclinação, respectivamente, no caso em que amostras de sistemas que não são “conhecidos” pela RNA_c são apresentados.

Deformações ocasionam alterações nos dados de entrada da rede estimadora RNA_E levando à degradação do desempenho e desvios nos valores de desalinhamento estimados. Uma segunda rede corretora RNA_C, treinada com amostras provenientes de diferentes sistemas constituídos de elementos com deformações aleatórias em suas superfícies, mostrou ser capaz de reduzir os erros das saídas da rede estimadora para novas amostras extraídas dos sistemas com deformações que foram apresentados à rede corretora durante treinamento. Para sistemas com deformações diferentes, porém de mesma magnitude, a estimativa do estado de alinhamento continua sendo afetada, mesmo após passar pela rede corretora.

Esse resultado indica que, quando as deformações superficiais nos elementos que compõem o sistema são conhecidas, é possível treinar uma rede corretora RNA_C dos dados estimados pela rede estimadora RNA_E, que reduza os erros de estimativa de desalinhamentos.

5.4 Sistema reflexivo com perturbações no espelho secundário e erros de posicionamento axial

Nesta seção, o mesmo sistema reflexivo analisado na seção 5.2 é agora analisado considerando-se também desvios de posicionamento axial.

5.4.1 Desalinhamentos no sistema

Para esta análise, deslocamentos axiais do segundo espelho em relação ao primeiro foram adicionados aos demais desalinhamentos analisados na seção 5.2. A Tabela 22 apresenta os casos analisados.

Tabela 22 – Descrição dos intervalos de desalinhamentos e amostragens adotadas na análise incluindo deslocamentos axiais.

	Desalinhamento	**Intervalo**	**Amostragem**
1	Sistema alinhado	-	-
2	DX	(-1,0 a 1,0) mm	0,01 mm
3	DY	(-1,0 a 1,0) mm	0,01 mm
4	TX	(-1,0 a 1,0)°	0,01°
5	TY	(-1,0 a 1,0)°	0,01°
6	AX	(-800,2 a -799,8) mm	0,01 mm
7	DX e DY	(-1,0 a 1,0) mm	0,02 mm
8	DX e TX	(-1,0 a 1,0) mm, (-1,0 a 1,0°)	0,02 mm e 0,02 °
9	TX e TY	(-1,0 a 1,0°)	0,02 °
10	DX e AX	(-1,0 a 1,0) mm, (-800,2 a -799,8) mm	0,02 mm e 0,01 mm
11	DY e AX	(-1,0 a 1,0) mm, (-800,2 a -799,8) mm	0,02 mm e 0,01 mm
12	TX e AX	(-1,0 a 1,0)°, (-800,2 a -799,8) mm	0,02 ° e 0,01 mm
13	TY e AX	(-1,0 a 1,0)°, (-800,2 a -799,8) mm	0,02 ° e 0,01 mm

Neste caso, além das 4 saídas já utilizadas no caso anterior, uma saída adicional foi usada para os deslocamentos axiais. Desta forma, a rede tem agora 5 saídas.

Novamente foram consideradas duas camadas escondidas e o número de neurônios em cada camada variou em diferentes treinamentos.

Deslocamentos axiais, em sistemas telescópicos com dois espelhos, induzem principalmente aberração esférica. Contudo, ainda que em menor escala, esses deslocamentos causam também coma. O efeito de erros de posicionamento axial no coma é, em geral, mais significativo em telescópios com dois espelhos aplanáticos. (5).

Dois treinamentos foram realizados, o primeiro considera como dados de entrada os coeficientes referentes à estratégia III, ou seja, Z_5, Z_6, Z_7 e Z_8 relacionados às aberrações coma e astigmatismo, medidos em 5 pontos ao longo do campo.

Um segundo treinamento foi realizado considerando-se agora um quinto coeficiente, o Z_9, que, no caso dos coeficientes de Zernike *Fringe*, refere-se à aberração esférica e *defocus* (44). A aberração esférica de terceira ordem, como dito anteriormente, não sofre influência de desalinhamentos como descentralização e inclinação, mas é influenciada por deslocamentos axiais e pode, portanto, ser importante no aprendizado da rede.

5.4.2 Resultados e discussões

Treinamento I:

Este treinamento foi realizado considerando-se dados de entrada de acordo com a estratégia III da Seção 5.2, num total de 20 tipos de dados de entrada. Os resultados apresentados a seguir correspondem a uma rede com 15 e 10 neurônios, respectivamente, nas camadas intermediárias e 5 neurônios na camada de saída. As camadas escondidas possuem função de ativação do tipo logística e a camada de saída possui função de ativação linear.

O gráfico de EQM em função da época de treinamento, para o primeiro treinamento realizado, é apresentado na Figura 66. A Tabela 23 apresenta o EQM para cada uma das saídas e o valor total diante das amostras de teste. O valor médio para cada saída e os desvios dos valores do EQM total em cada treinamento em relação ao seu valor médio também são apresentados. A Tabela 24 apresenta a porcentagem das amostras com erro, em módulo, inferior a 0,005mm ou 0,005°.

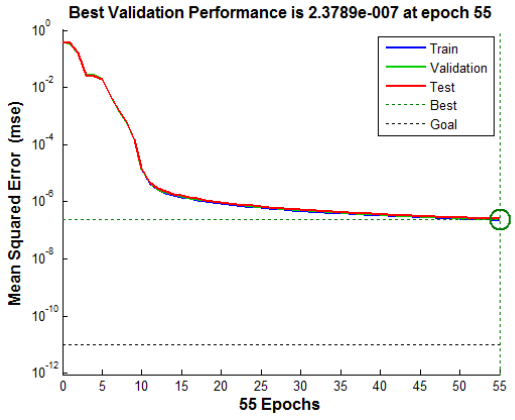

Figura 66 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada, treinamento conjunto 1.

Tabela 23 – EQM para cada uma das cinco saídas da RNA e valor total, media dos valores de cinco treinamentos e desvio dos valores de EQM total em relação ao seu valor médio, amostras referentes a estratégia III– amostras de teste.

Treinamento	**EQM**						**DESVIO EQM total**
	DX	DY	TX	TY	AX	TOTAL	
Conj. 1	1,2E-6	1,7E-6	7,2E-7	4,0E-7	3,9E-8	8,2E-7	-1,0E-9
Conj. 2	1,4E-6	1,0E-6	5,7E-7	3,0E-7	2,7E-8	6,7E-7	-1,5E-7
Conj. 3	2,7E-6	2,3E-6	1,2E-6	1,5E-6	2,2E-7	1,6E-6	7,8E-7
Conj. 4	3,6E-7	1,4E-6	3,1E-7	2,9E-7	1,5E-8	4,7E-7	-3,4E-7
Conj. 5	8,4E-7	1,2E-6	2,6E-7	3,1E-7	4,9E-8	5,2E-7	-2,9E-7
Valor Médio	1,3E-6	1,5E-6	6,1E-7	5,7E-7	7,0E-8	8,2E-7	

Tabela 24 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das cinco saídas da RNA, valor total e média, estratégia III – amostras de teste.

Treinamento	**AMOSTRAS C/ |ERRO|<0,005mm ou 0,005°**					
	DX	DY	TX	TY	AX	TOTAL
Conj. 1	99,5	99,2	99,8	99,9	100,0	99,7
Conj. 2	99,4	99,5	100,0	100,0	100,0	99,8
Conj. 3	98,5	98,4	99,6	99,5	100,0	99,2
Conj. 4	100,0	99,5	100,0	100,0	100,0	99,9
Conj. 5	99,9	99,3	100,0	100,0	100,0	99,9
Valor Médio	99,5	99,2	99,9	99,9	100,0	99,7

Histogramas de erros e gráficos de erro em função do número da amostra também são mostrados a seguir na Figura 67. A Figura 68 apresenta os histogramas de erros para cada uma das saídas, para as amostras de teste do primeiro conjunto treinado. Um total de 48.012 amostras foram extraídas deste sistema e 70% (33.608) delas foram usadas para treinamento, 15% (7.202) para validação e 15% para teste de desempenho.

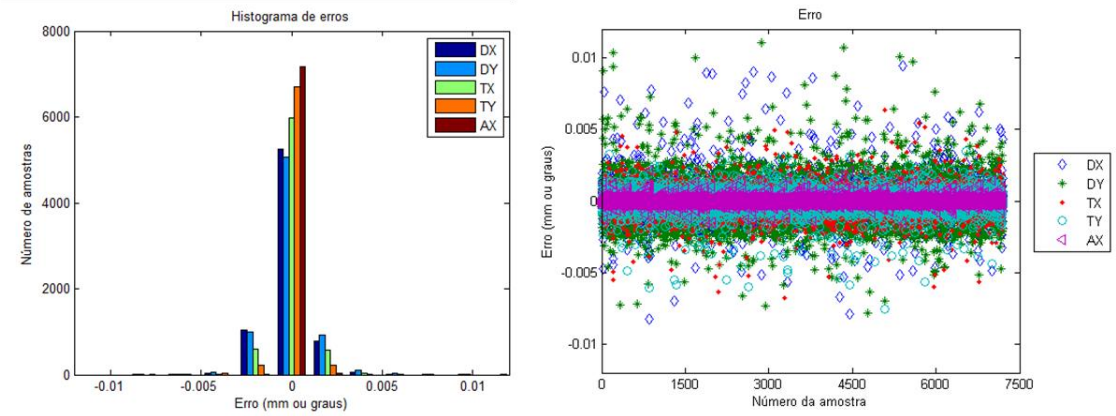

Figura 67 – Histograma de erros e gráfico de erro em função do número da amostra para as cinco saídas da rede treinada com as amostras referentes à estratégia III – treinamento 1, dados de teste.

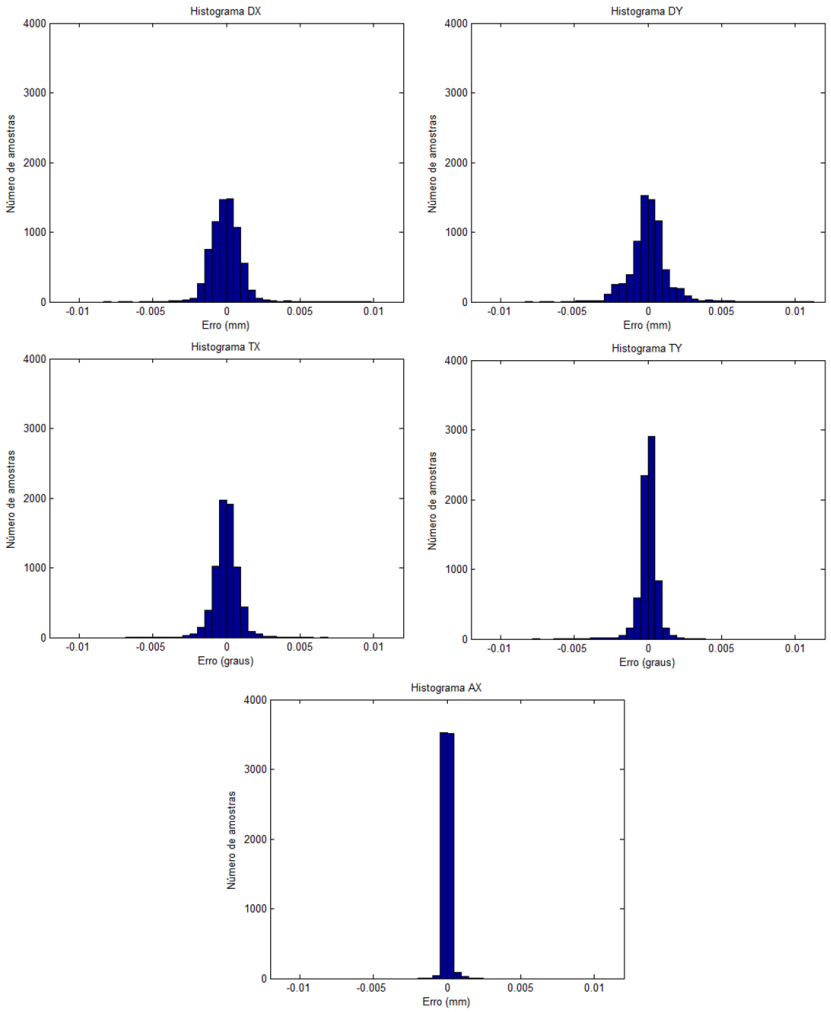

Figura 68 – Histogramas de erros para cada uma das cinco saídas para a RNA treinada com as amostras referentes à estratégia III, treinamento conjunto 1, dados de teste.

Treinamento II:

Este treinamento foi realizado considerando-se como dados de entrada os coeficientes de Zernike *FRINGE* Z_5, Z_6, Z_7, Z_8 e Z_9, medidos em cinco pontos ao longo do campo, totalizando 25 tipo de dados de entrada diferentes. Os resultados apresentados a seguir correspondem a uma rede com 15 e 10 neurônios, respectivamente, nas camadas intermediárias e 5 neurônios na camada de saída. As camadas escondidas possuem função de ativação do tipo logística e a camada de saída possui função de ativação linear.

Na Figura 69 é apresentado o gráfico de EQM em função da época de treinamento. A Tabela 25 apresenta os valores de EQM para cada saída da rede. A Tabela 26 apresenta a porcentagem de amostras de teste com |erro|<0,005 mm ou 0,005 °.

Histogramas de erros e gráficos de erro em função do número da amostra também são mostrados a seguir. Um total de 48.012 amostras foram extraídas deste sistema e 70% (33.608) delas foram usadas para treinamento, 15% (7.202) para validação e 15% para teste de desempenho.

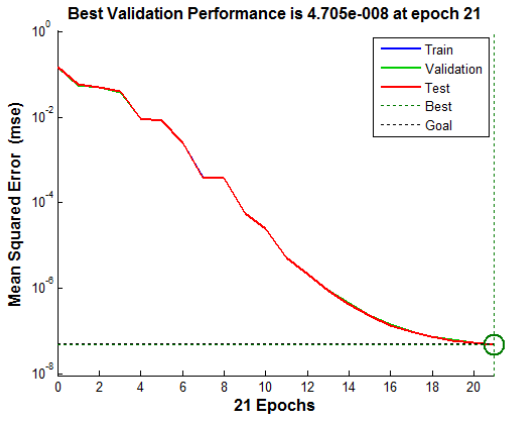

Figura 69 – Gráfico de EQM em função do número de épocas de treinamento para as amostras de treinamento, validação e teste da rede treinada, treinamento conjunto 1.

Tabela 25 – EQM para cada uma das cinco saídas da RNA e valor total, média dos valores de cinco treinamentos e desvio dos valores de EQM total em relação ao seu valor médio, amostras referentes aos treinamentos considerando Z_5, Z_6, Z_7, Z_8 e Z_9 medidos em cinco pontos no campo– amostras de teste.

Treinamento	EQM						DESVIO EQM total
	DX	DY	TX	TY	AX	TOTAL	
Conj. 1	3,1E-7	1,6E-7	1,6E-7	7,5E-8	8,9E-9	1,4E-7	-6,5E-7
Conj. 2	1,7E-6	1,5E-6	1,0E-6	6,9E-7	5,4E-8	9,9E-7	2,1E-7
Conj. 3	3,0E-7	4,6E-7	3,6E-7	4,4E-7	1,8E-8	3,1E-7	-4,8E-7
Conj. 4	1,4E-6	2,0E-6	2,1E-6	2,5E-6	1,6E-7	1,6E-6	8,4E-7
Conj. 5	1,8E-6	8,9E-7	7,6E-7	8,0E-7	4,2E-8	8,5E-7	6,7E-8
Valor Médio	1,1E-6	1,0E-6	8,8E-7	8,9E-7	5,1E-8	7,9E-7	

Tabela 26 – Porcentagem de amostras com |erro| < 0,005mm ou 0,005° para cada uma das cinco saídas da RNA, valor total e média, amostras referentes aos treinamentos considerando Z_5, Z_6, Z_7, Z_8 e Z_9 medidos em cinco pontos no campo– amostras de teste

Treinamento	AMOSTRAS C/ |ERRO|<0,005mm ou 0,005°					
	DX	DY	TX	TY	AX	TOTAL
Conj. 1	100,0	100,0	100,0	100,0	100,0	100,0
Conj. 2	99,0	99,2	99,8	99,9	100,0	99,6
Conj. 3	99,5	98,8	99,2	98,9	100,0	99,3
Conj. 4	98,9	99,7	99,7	99,9	100,0	99,6
Conj. 5	100,0	99,9	100,0	100,0	100,0	100,0
Valor Médio	99,5	99,5	99,7	99,7	100	99,7

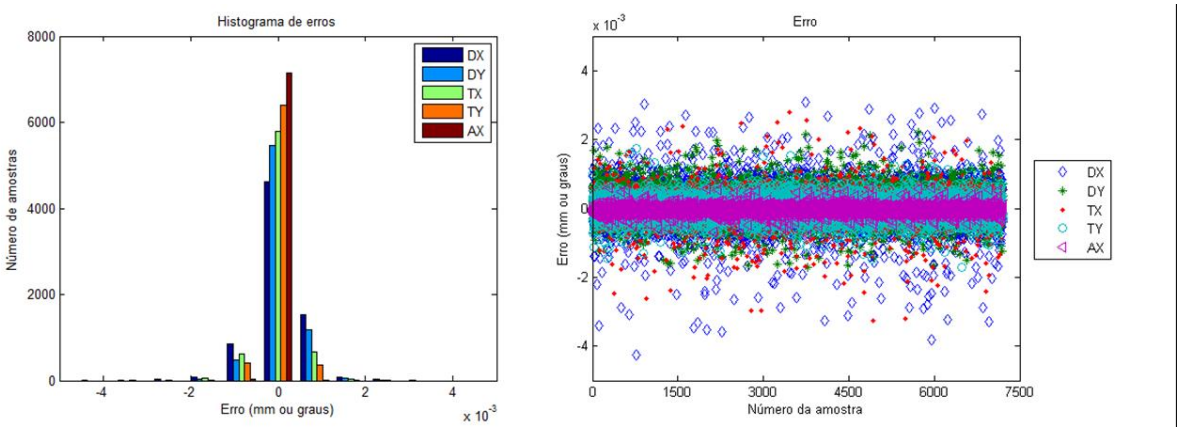

Figura 70 – Histograma de erros e gráfico de erro em função do número da amostra para as cinco saídas da rede treinada considerando coeficiente relativo à aberração esférica, treinamento conjunto 1 – amostras de teste.

A partir dos treinamentos realizados, observou-se uma ligeira melhora no desempenho da RNA, quando coeficientes referentes à aberração esférica foram incluidos nos dados de entrada. Contudo, os resultados obtidos, quando são apresentados apenas os coeficientes Z_5, Z_6, Z_7 e Z_8, são bons, com erros de até 0,012 mm ou graus. Os histogramas da Figura 71 indicam que a grande maioria das

amostras apresenta erros menores que 0,002 mm ou 0,002° no caso do treinamento do conjunto 1.

Ambos os conjuntos de dados de entrada resultaram em uma porcentagem média de 99,7% de amostras com erros inferiores a 0,005 mm ou 0,005° em cinco treinamentos.

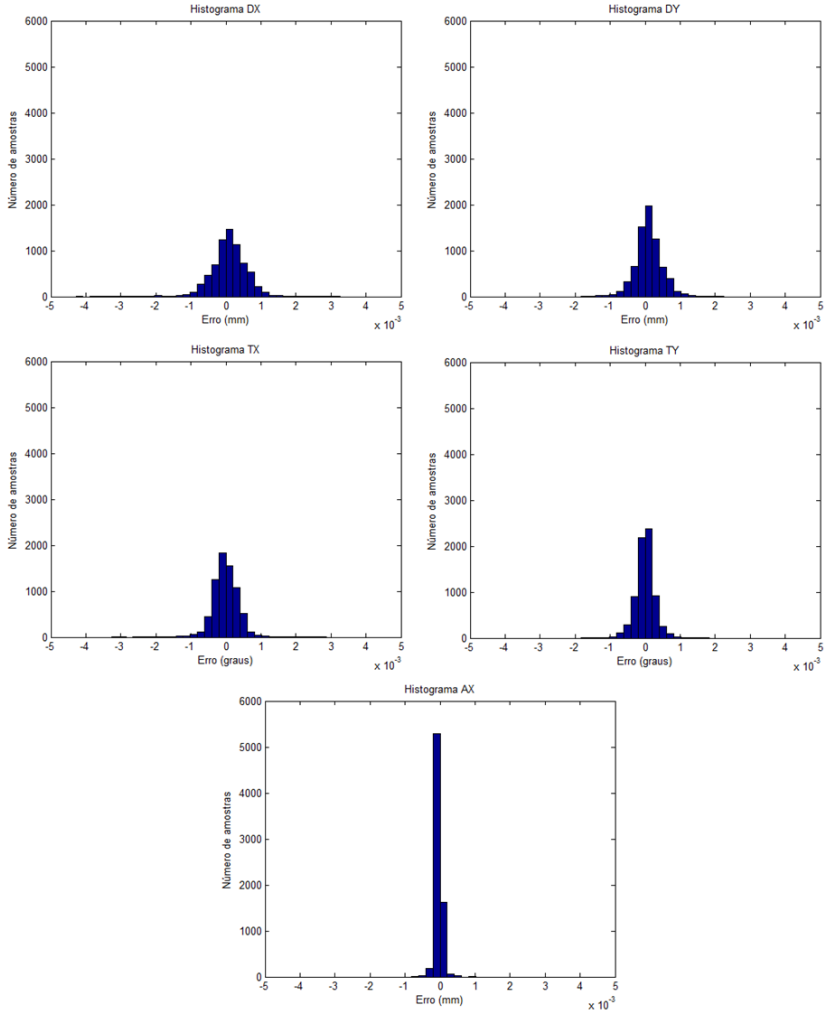

Figura 71 – Histogramas de erros para cada uma das cinco saídas para a RNA, treinamento conjunto 1.

Esse resultado indica que o conjunto de dados de entrada composto pelos coeficientes Z_5, Z_6, Z_7 e Z_8 extraídos em 5 pontos no campo contém informações suficientes, ao menos para o sistema analisado, para que a rede neural consiga identificar os deslocamentos axiais.

É importante lembrar que os termos Z_5 e Z_6, relacionados à astigmatismo, correspondem a astigmatismo e *defocus* (4, 44).

5.5 Sistema refrativo com vários elementos

5.5.1 Descrição do sistema

Nesta seção é apresentado um exemplo da aplicação do modelo proposto para o caso de um sistema refrativo composto por doze elementos.

O sistema, apresentado na Figura 72, possui as características descritas na Tabela 27 e tem aplicação em sensoriamento remoto orbital (76, 77).

A Figura 73 apresenta o gráfico da função de transferência de modulação[m] (MTF) que indica o seu desempenho na temperatura média de operação, considerando a faixa espectral de 450 a 890 nm.

Tabela 27 – Características do sistema óptico.

Descrição	Requisito
EFL (mm)	505,8
Abertura relativa	4,5
Campo de visada (graus)	8,8
Bandas espectrais (nm)	De 450 a 520, de 520 a 590, de 630 a 690, e de 770 a 890.
Resolução no solo (m)	20,0

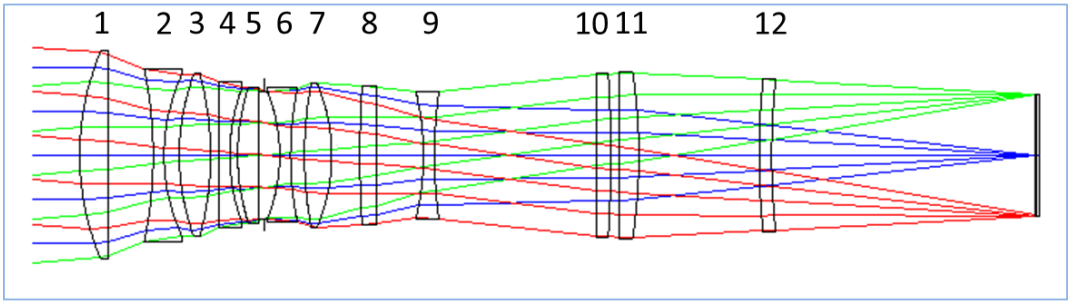

Figura 72 – Sistema óptico refrativo projetado para sensoriamento remoto orbital.

[m] MTF: Função de Transferência de Modulação representa uma medida da habilidade de um sistema óptico em transferir para a imagem os diferente níveis de detalhes do objeto. Fonte: The Photonics Dictionary, Pittsfield: Laurin Publishing, p. D89, 2004. Maiores informações podem ser encontradas em (17).

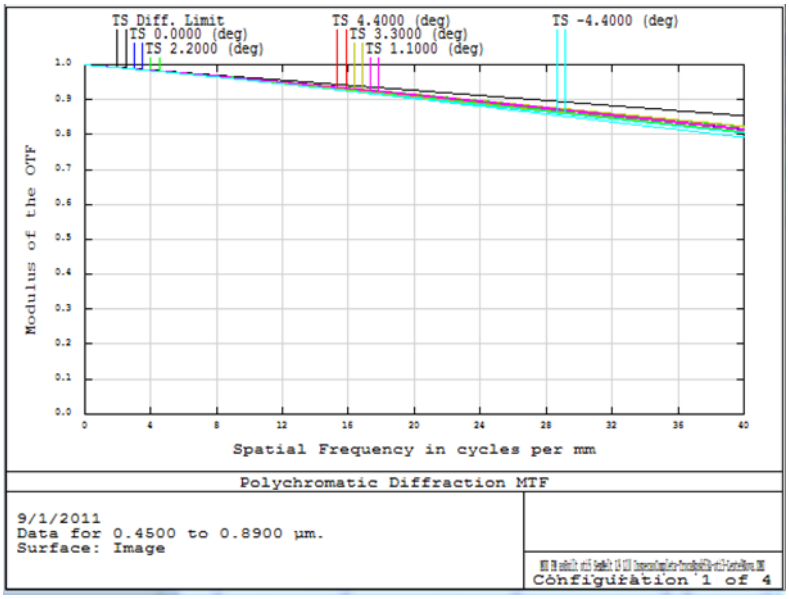

Figura 73 – Gráfico de MTF do sistema em toda faixa espectral.

Sempre que um sistema óptico é projetado, deve ser realizada uma análise de erros que determine os requisitos necessários para que, após fabricação e alinhamento, o sistema atinja as especificações do projeto. Uma análise completa engloba os erros de fabricação, deformações devido aos suportes mecânicos, erros de montagem, erros residuais relativos aos equipamentos de teste, efeitos ambientais, variações de temperatura, gradientes térmicos, vibração, etc.

As tolerâncias de montagem, obtidas a partir de uma análise de sensibilidade simplificada desse sistema, são apresentadas a seguir, na Tabela 28.

As tolerâncias consideradas, neste caso, correspondem a erros relativos aos valores conhecidos (erros de medida) dos raios de curvatura e espessura, que são diferentes das tolerâncias especificadas para fabricação, uma vez que o sistema passa por uma nova otimização *as built* após a fabricação de todos os componentes, e os espaçamentos entre lentes são reajustados de modo a compensar desvios dos valores finais após fabricação, em relação aos valores de projeto.

A análise foi feita baseada no critério de MTF, de modo que o desempenho final do sistema fosse aceitável, ou seja, acima do valor mínimo especificado de 0,65 @

38,5 lp/mm.

Tabela 28 – Tolerâncias de montagem e alinhamento do sistema – análise simplificada.

Descrição	Erro (±)
Descentralização L1 (*x* ou *y*) (mm)	0,010
Inclinação L1 (*x* ou *y*) (graus)	0,008
Descentralização L2 (*x* ou *y*) (mm)	0,005
Inclinação L2 (*x* ou *y*) (graus)	0,007
Descentralização L3 (*x* ou *y*) (mm)	0,005
Inclinação L3 (*x* ou *y*) (graus)	0,007
Descentralização L4 (*x* ou *y*) (mm)	0,010
Inclinação L4 (*x* ou *y*) (graus)	0,009
Descentralização L5 (*x* ou *y*) (mm)	0,010
Inclinação L5 (*x* ou *y*) (graus)	0,010
Descentralização L6 (*x* ou *y*) (mm)	0,005
Inclinação L6 (*x* ou *y*) (graus)	0,007
Descentralização L7 (*x* ou *y*) (mm)	0,005
Inclinação L7 (*x* ou *y*) (graus)	0,008
Descentralização L8 (*x* ou *y*) (mm)	0,020
Inclinação L8 (*x* ou *y*) (graus)	0,020
Descentralização L9 (*x* ou *y*) (mm)	0,020
Inclinação L9 (*x* ou *y*) (graus)	0,010
Descentralização L10 (*x* ou *y*) (mm)	0,100
Inclinação L10 (*x* ou *y*) (graus)	0,100
Descentralização L11 (*x* ou *y*) (mm)	0,090
Inclinação L11 (*x* ou *y*) (graus)	0,100
Descentralização L12 (*x* ou *y*) (mm)	0,100
Inclinação L12 (*x* ou *y*) (graus)	0,100
Espaçamento L1 – L2 (mm)	0,030
Espaçamento L2 – L3 (mm)	0,030
Espaçamento L3 – L4 (mm)	0,050
Espaçamento L4 – L5 (mm)	0,050
Espaçamento L5 – L6 (mm)	0,050
Espaçamento L6 – L7 (mm)	0,025
Espaçamento L7 – L8 (mm)	0,050
Espaçamento L8 – L9 (mm)	0,050
Espaçamento L9 – L10 (mm)	0,050
Espaçamento L10 – L11 (mm)	0,050
Espaçamento L11 – L12 (mm)	0,050

Além disso, por meio de uma simulação de Monte Carlo, o desempenho do sistema foi verificado. Nesta análise, um grande número de sistemas contendo erros dentro das tolerâncias especificadas são gerados de forma aleatória e o desempenho final do sistema é verificado. Desta forma, esta análise avalia diferentes possibilidades de acoplamentos de erros, inclusive não lineares, e fornece uma análise estatística do desempenho indicando as chances do sistema final, fabricado, atingir as especificações de projeto.

Os resultados apresentados na Tabela 28 referem-se às tolerâncias de montagem e alinhamento necessárias aos elementos individuais para que o sistema final tenha

desempenho satisfatório. Os valores apresentados na tabela, resultado dessa análise simplificada, não correspondem às tolerâncias finais do sistema que efetivamente foi construído.

No projeto original, diversos outros fatores foram considerados, conforme já mencionado anteriormente e, para que o sistema ficasse com tolerâncias mais amenas, possibilitando sua construção, algumas lentes, como é o caso das lentes 2 e 3, foram montadas em pequenos conjuntos independentes do sistema e alinhadas entre si. O mesmo ocorreu para outros conjuntos de lentes dentro do canhão. Desta forma, no alinhamento final do sistema como um todo, as lentes pré-alinhadas nesses "blocos de lentes" se movem juntas, tornando o sistema menos sensível a erros de posicionamento, e facilitando o alinhamento final. Além disso, compensadores foram usados para compensar erros de fabricação e alinhamento, um deles consiste no posicionamento axial do plano focal.

Não serão apresentados aqui mais detalhes sobre as soluções tomadas no projeto original, nem mesmo serão descritas as tolerâncias finais do projeto. Para a análise proposta, esses artifícios de projeto não precisam ser levados em consideração, uma vez que se pretende avaliar teoricamente o desempenho do estimador em encontrar os desalinhamentos no sistema. As tolerâncias de montagem deste sistema são apresentadas apenas para ilustrar a sua sensibilidade e indicar a precisão necessária de alinhamento. Dessa forma, as lentes são consideradas independentes entre si.

A Figura 74 ilustra estas diferenças de sensibilidade. Nela são apresentados gráficos da magnitude de astigmatismo (calculado pela equação $astigmatismo = \sqrt{Z_5^2 + Z_6^2}$) e coma (calculado pela equação $coma = \sqrt{Z_7^2 + Z_8^2}$), em função da posição no campo, para o sistema alinhado e para desalinhamentos do tipo descentralização e inclinação nas lentes 3, 5 e 7. Pode-se observar que, para este sistema, descentralizações influenciam mais o coma, e inclinações, o astigmatismo. As lentes 3 e 7 são mais sensíveis a desalinhamentos que a lente 5.

Na Figura 75 são apresentados os interferogramas do sistema, obtidos a partir do software Zemax, quando desalinhamentos de 0,1 mm ou 0,1 graus são inseridos nestes três elementos, mostrando, mais uma vez, a diferença de sensibilidade.

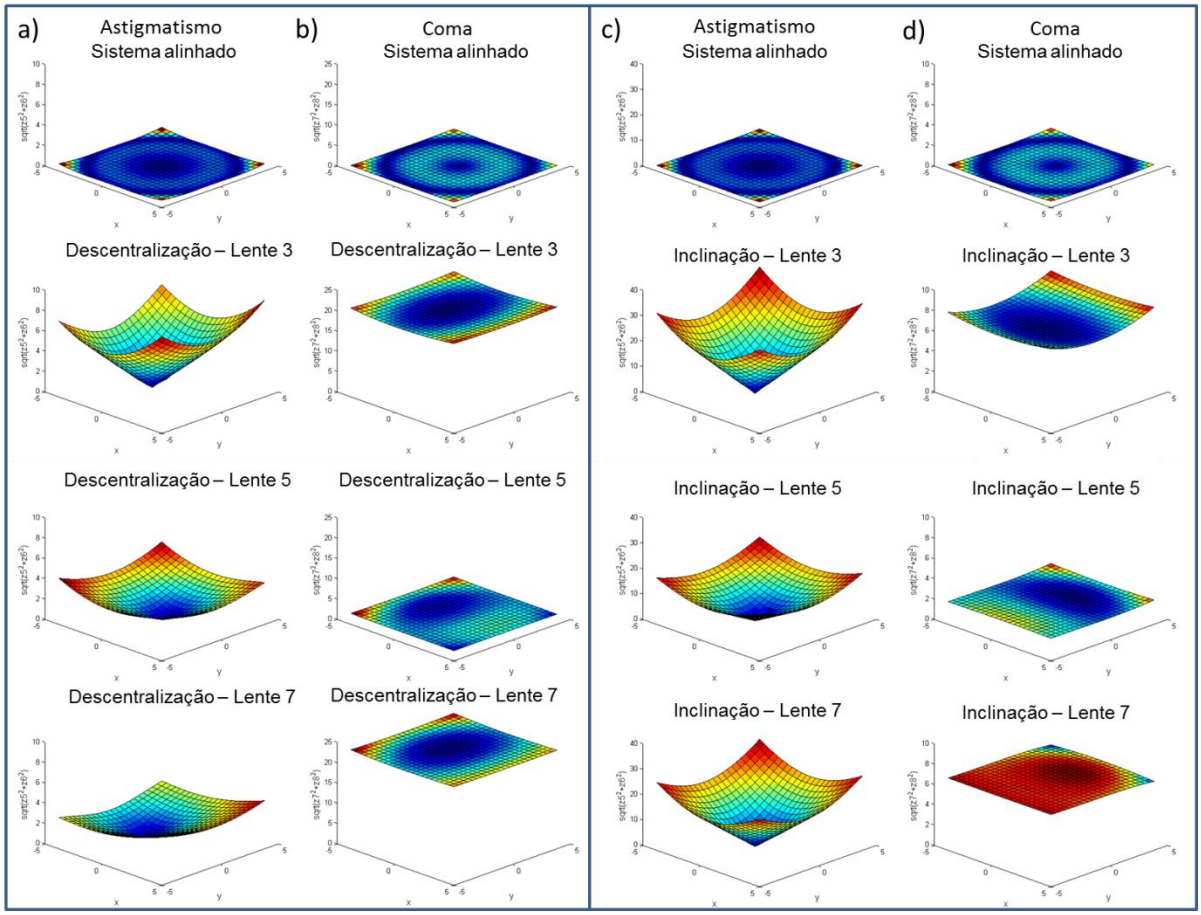

Figura 74 – Efeitos de desalinhamentos nas aberração do sistema; a) astigmatismo em todo o campo e b) coma em todo o campo, para o sistema alinhado e com desalinhamento do tipo descentralização nas lentes 3, 5 e 7, respectivamente, de cima para baixo; c) astigmatismo em todo o campo e d) coma em todo o campo, para o estado alinhado e para desalinhamentos do tipo inclinação nas lentes 3, 5 e 7.

5.5.2 Estimativa do estado de alinhamento

A análise realizada visa avaliar a capacidade de redes neurais artificiais em estimar o estado do alinhamento desse sistema, comprovando a eficiência da metodologia proposta, quando aplicada a sistemas ópticos complexos.

Nesta análise, três lentes (da esquerda para a direita na Figura 72, lentes: 3, 5 e 7) e dois espaçamentos entre lentes (entre lentes 3 e 4, e lentes 7 e 8) foram perturbados, considerando-se, inclusive, combinações de perturbações entre graus de liberdade de um mesmo elemento e entre os elementos, e resultando em 14

diferentes saídas (sendo quatro saídas referentes a DX, DY, TX e TY para cada lente e dois deslocamentos axiais AX).

Foram usados como dados de entrada os coeficientes Z_5, Z_6, Z_7, Z_8 e Z_9 extraídos em 5 posições ao longo do campo.

Um total de 43 combinações de perturbações foram consideradas e 53.823 amostras foram extraídas, sendo que 70% delas foram usadas para treinamento da rede, 15% para validação e 15% para teste do desempenho.

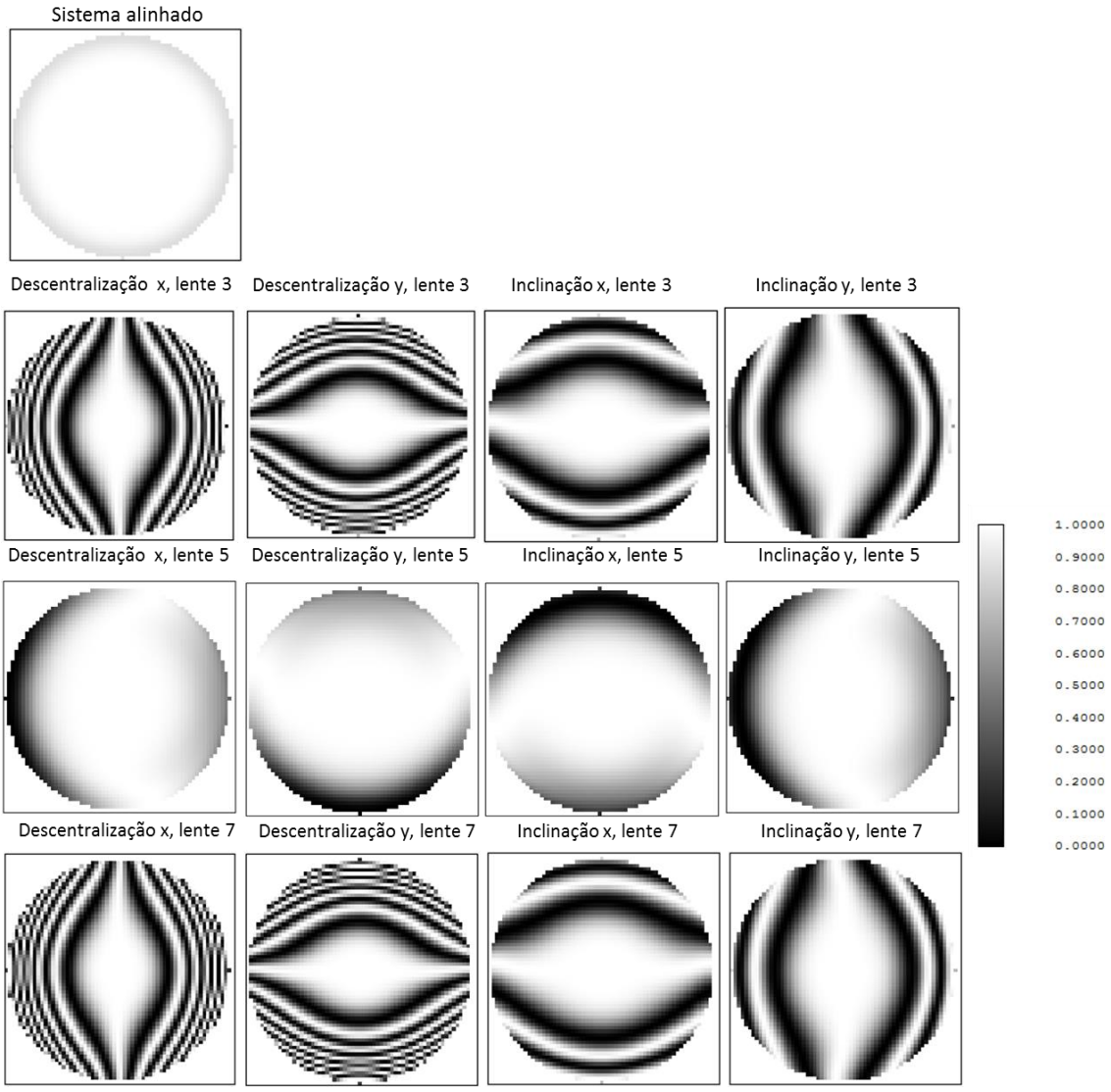

Figura 75- Interferogramas gerados no Zemax® para o sistema alinhado e para desalinhamentos do tipo descentralização e inclinação nas lentes 3, 5 e 7, respectivamente.

5.5.3 Resultados, discussões e conclusões

A RNA que apresentou melhor desempenho possui 25 e 42 neurônios, respectivamente, nas camadas intermediárias, com função de ativação do tipo logística. A camada de saída, com 14 neurônios, possui função de ativação linear. A Figura 76 apresenta o gráfico do EQM em função das épocas de treinamento.

Os histogramas de erros para os desalinhamentos são apresentados na Figura 77 a, para a lente 3, na Figura 77 b, para a lente 5, na Figura 77 c, para a lente 7 e na Figura 77 d, para os desalinhamentos axiais, treinamento 1.

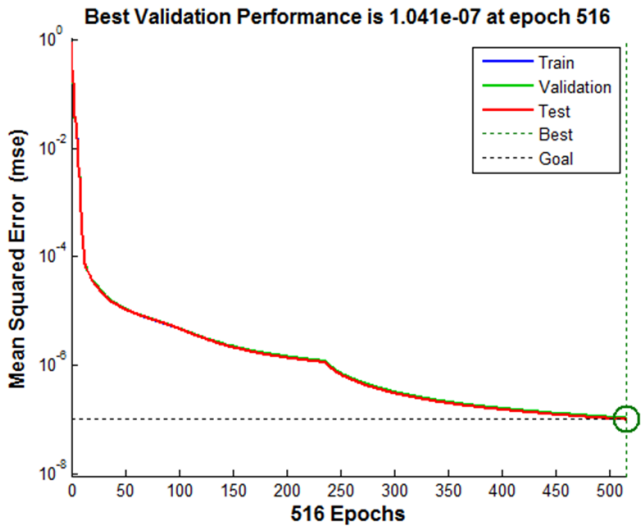

Figura 76 – EQM em função das épocas de treinamento, treinamento 1.

Os resultados mostram que 99,2% das amostras de teste resultaram em estimativas de desalinhamentos com erros menores que 0,0025 mm ou graus no caso do treinamento 1. O valor médio em cinco treinamentos foi de 98,6% (Tabela 29). Este resultado indica que a rede treinada possui sensibilidade suficiente para estimar desalinhamentos menores que os requisitados pelo sistema, apresentados na análise de tolerâncias (Tabela 28) mesmo no caso das lentes mais sensíveis.

Observa-se boa convergência do método quando aplicado a desalinhamentos múltiplos em sistemas complexos. Tal resultado indica que essa metodologia é promissora para as aplicações no alinhamento de tais sistemas.

O uso deste método pode significar ganhos no tempo do processo de alinhamento, que costumam ser longos e trabalhosos, e na qualidade final atingida após o alinhamento, já que a análise das aberrações presentes no sistema, através da medida da frente de onda, e utilizando redes neurais artificiais permite detectar desalinhamentos pequenos.

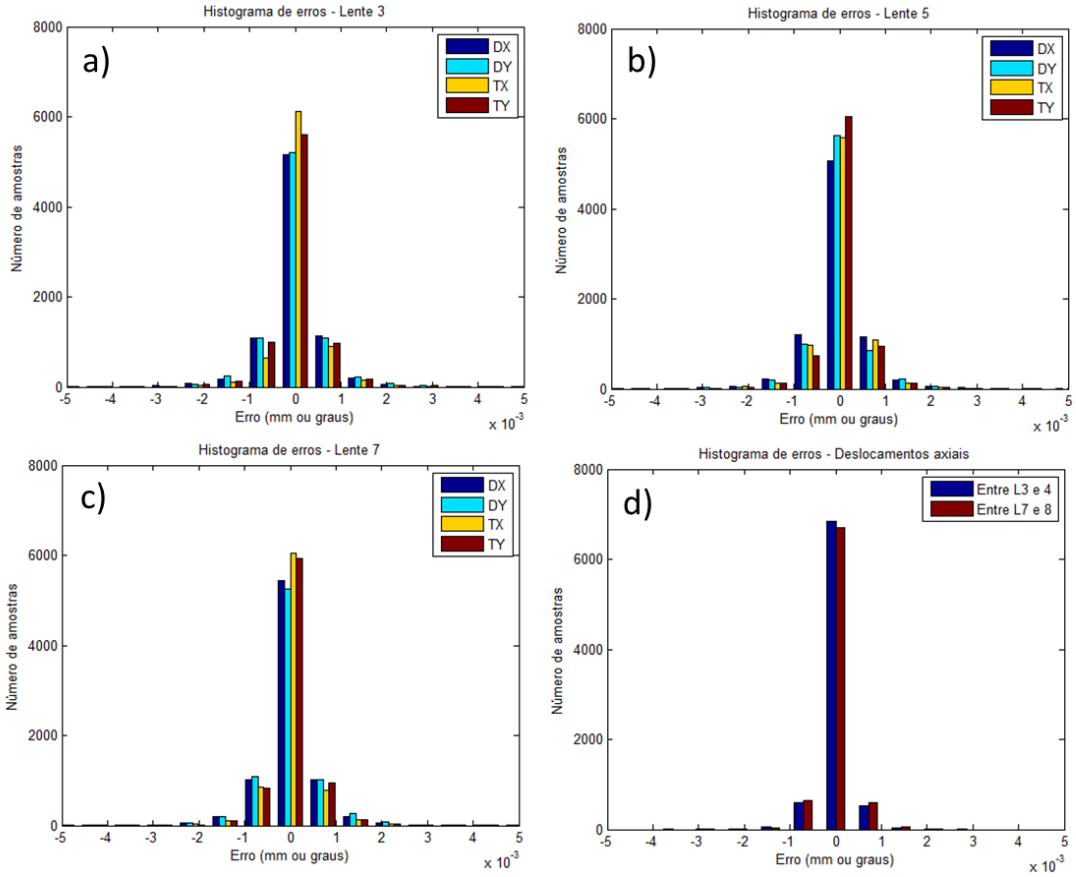

Figura 77 – Histogramas de erros para os desalinhamentos nas lentes a) 3, b) 5 e c) 7, e para d) deslocamentos axiais.

Tabela 29 – EQM em cada treinamento, valor médio em cinco treinamentos e desvio dos valores de EQM em relação ao seu valor médio. Porcentagem das amostras com erro, em módulo, inferior a 0,0025 mm ou 0,0025 °, valor médio e desvio do valor em cada treinamento em relação ao valor médio – amostras de teste.

	EQM	Desvio EQM	Amostras com |ERRO|<0,0025 (%)	Desvio amostras
Conj. 1	3,5E-7	-2,4E-7	99,2	0,6
Conj. 2	7,0E-7	1,2E-7	98,2	-0,4
Conj. 3	8,0E-7	2,2E-7	98,3	-0,3
Conj. 4	6,8E-7	9,4E-8	98,2	-0,4
Conj. 5	4,1E-7	-1,8E-7	99,0	0,4
Valor médio	5,9E-7		98,6	

6 CONCLUSÕES GERAIS E SUGESTÕES PARA TRABALHOS FUTUROS

6.1 Conclusões

O principal objetivo deste trabalho foi compreender os efeitos causados por desalinhamentos nas aberrações de sistemas ópticos e propor uma metodologia para alinhamento de sistemas ópticos baseada na análise dessas aberrações, a partir da frente de onda transmitida por esses sistemas.

Por esse motivo, a revisão bibliográfica abordou os tipos de aberrações monocromáticas presentes em sistemas rotacionalmente simétricos, em sistemas plano-simétricos e em sistemas sem simetria. Foram apresentadas a teoria de aberração de onda aplicável a sistemas plano-simétricos e a teoria de aberração de onda nodal, aplicável a sistemas com qualquer simetria. A teoria de aberração de onda nodal é mais ampla, porém, a teoria aplicável a sistemas plano-simétricos, por ser mais simples, é de mais fácil entendimento e capaz de prever os efeitos causados por desalinhamentos nesses sistemas.

Também foi apresentada uma breve revisão dos polinômios de Zernike, uma vez que as aberrações ópticas podem ser descritas por esses polinômios. Além disso, foram apresentadas, brevemente, as deformações superficiais em elementos ópticos e noções de interferometria, que pode ser usada na obtenção da frente de onda transmitida em sistemas reais.

A revisão bibliográfica abordou também a teoria básica de redes neurais artificiais, adotadas neste trabalho para a estimativa de desalinhamentos em sistemas ópticos.

Neste trabalho, foi apresentado um estudo da sensibilidade ao alinhamento de sistemas ópticos reflexivos compostos por dois elementos. Diversas configurações de telescópios do tipo Cassegrain, diferentes entre si apenas devido às constantes cônicas dos espelhos, foram analisadas no que diz respeito à sua sensibilidade à inclinação e descentralização do espelho secundário. Dentre as configurações

consideradas, uma específica apresentou menor sensibilidade à desalinhamentos do tipo descentralização. Para entender melhor este comportamento, aplicou-se a teoria de aberração de onda para sistemas plano-simétricos, proposta por Sasián e apresentada com detalhes no Capítulo 2, e constatou-se que a escolha apropriada da constante cônica do espelho secundário faz com que coma uniforme de terceira ordem seja compensado quando descentralizações estão presentes nesse espelho, fazendo com que este sistema seja livre da aberração mais importante causada por desalinhamentos neste tipo de telescópio.

Foi também proposta uma metodologia de alinhamento baseada na análise da frente de onda transmitida por sistemas ópticos, que utiliza redes neurais artificiais para a estimativa dos erros de alinhamento. Redes neurais artificiais foram empregadas na análise dos coeficientes dos polinômios de Zernike para estimar desalinhamentos tanto em sistemas reflexivos como em sistemas refrativos. As principais conclusões obtidas no que diz respeito à metodologia proposta foram:

- o emprego de redes neurais artificiais na estimativa de desalinhamentos em um sistema reflexivo composto por dois elementos, perfeito (sem considerar deformações superficiais nos elementos e sem ruído nos dados), foi satisfatório para os quatro tipos de conjunto de dados de entrada considerados. Foram obtidos bons resultados do estimador mesmo quando apenas os quatro tipos de coeficientes de Zernike relativos a coma e astigmatismo foram extraídos apenas sobre o eixo óptico. Resultados melhores foram obtidos quando medidas em outros quatro pontos fora do eixo foram consideradas. Também ocorreu uma ligeira melhora de desempenho quando coeficientes relacionados a inclinação e a *defocus* foram considerados na medida, contudo, o uso de coeficientes relativos apenas a coma e astigmatismo são suficientes para a estimativa de desalinhamentos dos tipos descentralização e inclinação;

- redes neurais artificiais também apresentaram desempenho satisfatório quanto à estimativa de erros de posicionamento axial dos elementos. Uma análise considerando os mesmo quatro coeficientes dos polinômios de Zernike relativos a coma e astigmatismo e uma segunda análise considerando-se também o coeficiente relativo à aberração esférica foram apresentadas e ambas apresentaram bons resultados. Uma ligeira melhora foi observada quando o coeficiente relativo à aberração esférica foi usado;

- deformações superficiais ocasionam degradação de desempenho do estimador. Uma segunda rede neural “corretora” dos dados estimados, pode ser usada e melhora ligeiramente os valores estimados, reduzindo o erro de estimativa causado pelas deformações quando estas são “conhecidas” pela rede. A rede corretora não reduziu o erro de estimativa para amostras de sistemas com deformações diferentes dos apresentados à rede corretora durante treinamento;

- análises realizadas em sistemas refrativos mostraram que esta metodologia baseada no uso de redes neurais artificiais também é capaz de estimar desalinhamentos nesses sistemas;

- a aplicação desse método, de forma simplificada, em um sistema refrativo composto por vários elementos, indicou que essa metodologia é promissora quanto ao seu uso em sistemas mais complexos.

Pode-se concluir que os estudos realizados mostraram que redes neurais artificiais, quando aplicadas na análise e estimativa de desalinhamentos, fornecem resultados satisfatórios quanto à estimativa de desalinhamentos dos tipos descentralização, inclinação e erros de posicionamento axial. Além disso, quando são considerados erros de superfície, que ocorrem em sistemas reais, redes neurais artificiais ainda podem fornecer indicações dos desalinhamentos presentes no sistema, quando essas deformações são conhecidas. Acredita-se que redes neurais artificiais constituem uma alternativa promissora no alinhamento de sistemas ópticos complexos.

6.2 Sugestões para trabalhos futuros

Baseado nos resultados obtidos neste trabalho, algumas sugestões podem ser feitas quanto à sua continuação.

- O trabalho apresentado consistiu em uma análise teórica do uso de redes neurais artificiais como estimadores de desalinhamentos em sistemas ópticos. Seria interessante que a metodologia proposta fosse aplicada a um sistema real, e que

fossem avaliados também os efeitos que os erros de medida e as condições ambientais ocasionam no desempenho do estimador.

- Considerando-se os resultados promissores obtidos teoricamente, quando redes neurais artificiais do tipo Perceptron Multicamadas são usadas para a estimativa de desalinhamentos em sistemas ópticos, seria interessante avaliar também o desempenho de outros tipos de redes neurais, como por exemplo, as redes de funções de base radial (RBF).

- Na metodologia proposta nesta tese, foram consideradas diferentes saídas de um mesmo MLP para os diferentes graus de liberdade do sistema. Uma abordagem alternativa, seria o uso de uma rede com uma única saída para cada grau de liberdade.

- Seria interessante também usar ferramentas de processamento de sinais como a chamada PCA (Análise de Componentes Principais) para dar às variáveis mais representatividade e reduzir a dimensionalidade da rede.

- Mais estudos devem ser realizados buscando encontrar alternativas de tornar estimadores de desalinhamentos baseados em redes neurais artificiais mais robustos quanto a erros de formas nas superfícies dos componentes do sistema.

REFERÊNCIAS

1 SHANNON, R. R**. The art and science of optical design**. New York: Cambridge University Press, 1997. 614p.

2 THOMPSON, K. P. **Aberration fields in tilted and decentered optical systems**. 1980. 273p. Tese (Doutorado em óptica) - The University of Arizona, Tucson-Arizona, 1980.

3 SASIÁN, J. M. How to approach the design of a bilateral symmetric optical system. **Optical Engineering**, v. 33, n. 6, p. 2045-2061, 1994.

4 MAHAJAN V. N. Zernike circle polynomial and optical aberrations of systems with circular pupils. **Applied Optics**, v. 33, n. 34, p. 8121-8124, 1994.

5 WILSON R.N., **Reflecting telescope optics I**: Basic design theory and its historical development. 2nd ed. New York: Springer, 2000. 543p.

6 MANUEL, A. M. **Field-dependent aberrations for misaligned reflective optical systems**. 2009. 404p. Tese (Doutorado em óptica) - College of Optical Sciences – The University of Arizona, Tucson – Arizona, 2009.

7 SCHMID, T. **Misalignment induced nodal aberration fields and their use in the alignment of astronomical telescopes**. 2010. 270p. Tese (Doutorado em óptica) - College of Optics and Photonics - University of Central Florida, Orlando - Florida, 2010.

8 CODE V enews – Quick tip – finding the most effective compensators during tolerancing and alignment. Optical Research Associates. 2007. Disponível em: <http://www.opticalres.com/E-News/cvnews07/cvnews0807.html>. Acesso em: 14 abr. 2013.

9 ANGEL, J. R. P.; WIZINOWICH, P.; LLOYD-HART, M.; SANDLER, D. Adaptive optics for array telescopes using neural-network techniques. **Nature**, v. 348, p. 221-224, 1990. doi:10.1038/348221a0

10 SANDLER, D. G.; BARRETT, T. K.; FUGATE, R. Q. Recovery of atmospheric phase distortion from stellar images using an artificial neural network. In: EALEY, M.

A.(Ed.). **Active and adaptive optical components**. San Diego: SPIE,1992. p.491-499.(Proceedings of SPIE, v. 1543).

11 SANDLER, D. G.; BARRETT, D. A.; PALMER, D. A.; FUGATE, R. Q.; WILD, W. J. Use of neural network to control an adaptive optics system for an astronomical telescope. **Nature**, v. 351, p.300-302, 1991.

12 BARRETT, T. K.; SANDLER, D. G. Artificial neural network for the determination of Hubble Space Telescope aberration from stellar images. **Applied Optics,** v. 32, n. 10, p.1720-1720, 1993.

13 HART, G. A.; BAILEY, A. W.; PALUMBO, L. J. Alinet: neural net automatic alignment of high energy laser resonator optical elements. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, IJCNN, 1992, Baltimore, MD. **Proceedings…** Baltimore: IEEE,1992. v. 2I, p.823-83.

14 YOON, H. K.; GWEON, D. G.; LEE, J. H.; JEONG, J.; OH, H. R. A neural-network approach based on interference pattern analysis: application to an autoalignment method for the focusing unit of NFR system. **Japanese Journal of Applied Physics**, v. 43, n. 7B, p.4790-4794, 2004.

15 HAYKIN, S. **Redes neurais:** princípios e prática. Trad. ENGEL, P. M. 2nd ed. Porto Alegre: Bookman, 2001. 900p.

16 CODE V enews – CODE V tip - using CODE V's alignment feature to respace a lens. Optical Research Associates. 2008. Disponível em: <http://www.opticalres.com/E-News/cvnews07/cvnews0807.html>. Acesso em: 14 abr. 2013.

17 SMITH, W. J. **Modern optical engineering**, 4th ed. New York: McGraw Hill, 2008. 754p.

18 RIMMER, M. Analysis of perturbed lens systems. **Applied Optics,** v. 9, n.3, p.533-538, 1970.

19 BUCHROEDER, R. A. **Tilted component optical systems**. 1976, 173p. Tese (Doutorado em óptica) - The University of Arizona, Tucson - Arizona, 1976.

20 SCHMID, T.; THOMPSON, K.; ROLLAND, J. Misalignment-induced nodal aberration fields in two-mirror astronomical telescopes. **Applied Optics**, v. 49 n.16, p.D131-144, 2010.

21 WILSON R.N., **Reflecting telescope optics II**. 2nd ed. New York: Springer, 2001.554p.

22 SCHMID, T.; THOMPSON, K. P.; ROLLAND, J. P. A unique astigmatic nodal property in misaligned Ritchey-Chrétien telescopes with misalignment coma removed. **Optics Express,** v. 18, n. 5, p.5282-5288, 2010.

23 THOMPSON, K. P.; SCHMID, T.; CAKMAKCI, O.; ROLLAND, J. P. Real-ray-based method for locating individual surface aberration field centers in imaging optical systems without rotational symmetry. **Journal of the Optical Society of America A,** v. 26 n. 6, p.1503-1517, 2009.

24 YANG, X.; HAN, C.; YU, J. Method for computer-aided alignment of complex optical system. In:INTERNATIONAL SYMPOSIUM ON ADVANCED OPTICAL MANUFACTURING AND TESTING TECHNOLOGIES: OPTICAL TEST AND MEASUREMENT TECNOLOGY AND EQUIPMENT,2, Xian, China, 2006. **Proceedings…**Xian: SPIE, 2006**.** v.6150, p.61503X1-7.

25 LUNDGREN, M. A.; WOLFE, W. L. Alignment of a three-mirror off-axis telescope by reverse-optimization. **Optical Engineering,** v. 30, n. 3, p.307-311, 1991.

26 JEONG, H. J.; LAWRENCE, G. N; NAHM, K. B. Auto-alignment of a three-mirror off-axis telescope by reverse optimization and end-to-end aberration measurements. In: CURRENT DEVELOPMENTS IN OPTICAL ENGINEERING,2,I, San Diego, CA, 1986. **Proceedings…** San Diego**:** SPIE, 1987.v. 0818, p.419.

27 LUNDGREN, M. A.; WOLFE, W. L. Simultaneous alignment and multiple surface figure testing of optical system components via wavefront aberration measurement and reverse optimization. In: INTEL. LENS DESIGN CONF, 1990, Monterey, USA, 1991.**Proceedings…** Monterey: SPIE, 1991, v. 1354 p.533-539.

28 HUANG, Y.; LI, L.; CAO, Y. Computer-aided alignment for space telescope optical system. In: INTERNATIONAL SYMPOSIUM ON ADVANCED OPTICAL MANUFACTURING AND TESTING TECHNOLOGIES: ADVANCED OPTICAL MANUFACTURING TECNOLOGIES, 2, Xian, China, 2005. **Proceedings…** Xian:SPIE, 2006. v. 6149, p.61490P-1-7.

29 ZHANG, B.; ZHANG, X.; WANG, C.; HAN, C. Investigation on computer-aided alignment of the complex optical system. In: ADVANCED OPTICAL MANUFACTURING AND TESTING TECHNOLOGY, Chengdu, China, 2000. **Proceedings…** Chengdu: SPIE, 2000. v. 4231, p.67-71.

30 ALBUQUERQUE, B. F. C.; LOPES, R. V. F.; KUGA, H. K.; CARVALHO, E. G.; SCADUTO, L. C. N.; et al. Misalignment parameters estimation in refractive optical systems. In: OPTICAL SYSTEM ALIGNMENT AND TOLERANCING, 2, San Diego, CA, USA, 2008. **Proceedings…** San Diego: SPIE, 2008. v. 70680P. doi:10.1117/12.797711.

31 LEE, H.; DALTON, G.; TOSH, I.; KIM, S. Computer-aided alignment I: Phase and amplitude modulation of alignment-influenced optical wavefront. **Optics Express,** v. 15, n. 6, p.3127-3139, 2007.

32 LEE, H.; DALTON, G.; TOSH, I.; KIM, S. Computer-aided alignment II: Optical system alignment using differential wavefront sampling. **Optics Express,** v. 15, n. 23, p.15424-15436, 2007.

33 LEE, H.; DALTON, G.; TOSH, I.; KIM, S. Computer-aided alignment III: Description of inter-element alignment effect in circular-pupil optical systems. **Optics Express,** v. 16, n. 15, p.10992-11006, 2008.

34 TESSIERES R. **Analysis for alignment of optical systems**. 2003, 177p. Dissertação (Mestrado em óptica) - The University of Arizona, Tucson – Arizona, 2003.

35 AGOSTINHO, O. L.; RODRIGUES, A. C. S.; LIRANI, J. **Tolerâncias, ajustes, desvios e análise de dimensões.** São Paulo: Editora Edgard Blucher Ltda. 2001, 296p.

36 MAHAJAN V. N. **Optical imaging and aberrations - part I:** ray geometrical optics, Bellingham: SPIE PRESS, 1998. 470p.

37 SASIÁN, J. M. Theory of sixth-order wave aberrations, **Applied Optics**, v. 49, n.16, p.D69-D95, 2010.

38 HOPKINS, G. W. Proximate ray tracing and wave aberration coefficients. **Journal of the Optical Society of America,** v. 66, n. 9, p.942-949, 1976.

39 GEARY J. M. **Introduction to lens design:** with practical ZEMAX® examples. Richmond: Willmann-Bell, Inc., 2007. 462 p.

40 SLYUSAREV, G. G. **Aberration and optical design theory**. 2nd ed. Redeliffe Way: Adam Hilger., 1984. 660p.

41 SCHMID, T.; THOMPSON, K.; ROLLAND, J. Alignment of two mirror astronomical telescopes: the astigmatic component. In: MODELING, SYSTEMS ENGINEERING,AND PROJECT MANAGEMENT FOR ASTRONOMY, III, Marseille, France, 2008. **Proceedings…** Marseille, France: SPIE, 2008. v. 7017, p.70170C.

42 REN, B.; JIN, G.; ZHONG, X. Third-order coma-free point in two-mirror telescopes by a vector approach. **Applied Optics**, v. 50 n. 21, p.3918-3923, 2011.

43 MAHAJAN V. N. Zernike polynomial and optical aberrations. **Applied Optics**, v. 34, n.34, p.8060-62, 1995.

44 WYANT, J. C. E CREATH, K. Basic wavefront aberration theory for optical metrology. In: SHANNON R. R., WYANT J. C. **Applied optics and optical engineering**. New York: Academic Press, 1992. v.11, cap. 1, p. 27-39. Disponível em: <http://www.optics.arizona.edu/jcwyant/zernikes/Zernikes.pdf>. Acesso em: 15 out. 2011.

45 ZEMAX Optical design program user's manual. Tucson, USA: Radiant ZEMAX LCC, 2011. 805p.

46 MAHAJAN V. N. **Optical imaging and aberrations,part II:** wave diffraction optics, Bellingham: SPIE PRESS, 2001. 463p.

47 OPTICAL RESEARCH ASSOCIATES. **Appendix C: Zernike polynomial - CODE V 10.2 reference manua**l, p.C1-C10, 2011. Disponível em: <http://www.cadfamily.com/download/Optical/CodeV-Wav/appendixc.pdf>. Acesso em: 17 out. 2011.

48 ZYGO CORPORATION. **Zernike MetroPro Application – OMP-0383B**, Middlefield: Zygo Corporation, 2002, 12p.

49 ZERNIKE Polynomials. Tucson, AZ. Disponível em: <http://wyant.optics.arizona.edu/zernikes/zernikes.htm>. Acesso em: 20 mai. 2013.

50 ZERNIKE Polynomials. Tucson, AZ. Disponível em: <http://fp.optics.arizona.edu/jcwyant/Zernikes/ZernikePolynomials.htm>. Acesso em: 20 mai. 2013.

51 GRAY, R. W.; DUNN, C.; THOMPSON, K. P.; ROLLAND, J. P. An analytic expression for the field dependence of Zernike polynomials in rotationally symmetric optical systems. **Optics Express**, v. 20, n. 15 p.16436-16449, 2012.

52 AIKENS, D. M.; DEGROOTE, J. E.; YOUNGWORTH, R. N., **Specification and control of mid-spatial frequency wavefront errors in optical systems**. Disponível em: <http://www.opticsinfobase.org/abstract.cfm?URI=OFT-2008-OTuA1>. Acesso em: 20 ago. 2013.

53 PORTO, A. J. V. et al. **Usinagem de ultraprecisão**. São Carlos: RiMa, FAPESP, 2004. 293p.

54 KOTHA, A.; HARVEY, J. E. Scattering effects of machined optical surfaces. In: OPTICAL SCATTERING IN THE OPTICS, SEMICONDUCTOR, AND COMPUTER DISK INDUSTRIES, San Diego, CA, USA, 1995. **Proceedings…** San Diego, USA: SPIE, 1995. v. 2541, p. 54-65.

55 CHURCH, E. L.; TAKACS, P. Z. Specification of surface figure and finish in terms of system performance. **Applied Optics,** v. 32, n19, p.3344-3353, 1993.

56 YOUNGWORTH, R. N.; STONE, B. D. Simple estimates for the effects of mid-spatial-frequency surface errors on image quality. **Applied Optics,** v. 39, n. 13, p.2198-2209, 2000.

57 HARVEY, J. E.; LEWOTSKY, K. L.; KOTHA, A. Effects of surface scatter on the optical performance of x-ray synchrotron beam-line mirrors. **Applied Optics,** v. 34, n. 16, p.3024-3032, 1995.

58 EVANS, C. J.; PARKS, R. E.; SULLIVAN, P. J.; TAYLOR, J. S. Visualization of surface figure by the use of Zernike polynomials. **Applied Optics**, v. 34 n. 34, p.7815-7819, 1995.

59 KIM, N. H. **How to use TEZI to tolerance for manufacturing-related surface sag error**. 2005. Disponível em <http://www.radiantzemax.com/kb-en/KnowledgebaseArticle50208.aspx> . Acesso em: 27 fev. 2013.

60 ZYGO CORPORATION. **Interferogram interpretation and evaluation handbook**. 4th. ed. Middlefield: ZYGO Corporation, 1993. 23p.

61 GEARY, J. **Introduction to Optical Testing**. Bellingham: SPIE Press, 1993. 149p.

62 ZILIO, S. C. **Optica moderna:** fundamentos e aplicações. São Carlos: IFSC-USP. 2009. 306p.

63 GOODWIN, E. P.; WYANT, J. C. **Field guide to interferometric optical testing**. v. FG10. Bellingham: SPIE Press, 2006.114p.

64 SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais:** para engenharia e ciências aplicadas. São Paulo: Artliber, 2010. 399 p.

65 DEMUTH, H.; BEALE, M. **Neural network toolbox:** for use with matlab – user's guide, version 4, Natick, MA: The MathWorks, Inc., 2002.

66 BEALE, M. H.; HAGAN, M. T.; DEMUTH, H. B. **Neural network toolbox – user's guide R2012b**. Natick, MA: The MathWorks, Inc., 2012.

67 SCADUTO, L. C. N.; SASIÁN, J.; STEFANI, M. A.; CASTRO NETO, J. C. Two-mirror telescope design with third-order coma insensitive to decenter misalignment. **Optics Express**, v. 21, n. 6, p.6851-6865, 2013.

68 SASIÁN, J. M. The world of unobstructed reflecting telescopes. **ATM Journal**, v. 1, p. 10-15, 1992.

69 MCLEOD B. A., Collimation of fast wide-field telescopes. **Publications of the Astronomical Society of the Pacific**, v. 108, p.217-219, 1996.

70 DUBIN, M.; LAMPEN, S.; BURGE, J. Characterization of alignment using measurements of the sine condition. In: OPTICAL SYSTEM ALIGNMENT, TOLERANCING, AND VERIFICATION, V, San Diego, CA, USA, 2011. **Proceedings…** San Diego, CA, USA: SPIE, 2011. v. 8131 p.813107-1-9.

71 MZDDE – The ZEMAX DDE toolbox for Matlab. Disponível em: < file:///C:/Users/Public/Documents/MZDDE/Documents/ZEMAX%20DDE%20Toolbox%20for%20Matlab.html>. Acesso em: 10 ago. 2013.

72 SOLA, J. E; SEVILLA, J. Importance of input data normalization for the application of neural networks to complex industrial problems. **IEEE Transactions on Nuclear Sciences**, v. 44, n. 33, p.1464-1468, 1997.

73 GONZAGA A., **Métodos de avaliação de classificadores** – aula 9. São Carlos: EESC-USP, s.d. Notas de aula.

74 ZYGO CORPORATION. **Zygo's guide to typical interferometer setups**. 4 version. 2003. Disponível em < http://www.zygo.com/met/setupsguide/zygo_int_setups.pdf>. Acesso em: 10 ago. 2013.

75 ZYGO CORPORATION. **FAQ: What is the correct Interferometric Scale Factor (ISF)?**. Disponível em: <http://www.zygo.com/?sup=/resource/faqs.cgi?id=60.>. Acesso em: 21 mar. 2013.

76 CARVALHO, E. G.; SCADUTO, L. C. N.; SOARES, A. L.; SANTOS, A. G.; EVANGELISTA, S. H.; REBOLHO, D.; SANTOS, A. R.; CASTELLAR, A.; AZEKA, L. A. L.; SOUZA, W. M.; MODUGNO, R. G.; CARTOLANO, R.; VALES, L. F.; MALAVOLTA, A. T.; BARBALHO, S.; SANTOS, F. S.; CANDELORO, L.; SEGORIA, D.; STEFANI, M. A.; CASTRO NETO, J. C. The Brazilian Multispectral Camera (MUX) for the China_Brazil Earth Resources Satellite - CBERS 3 & 4. In: INTERNATIONAL CONFERENCE ON SPACE OPTICS, ICSO, 2010, Rhodes Island, Greece. **Proceedings**...Rhodes Island: ICSO 2010.

77 SCADUTO, L. C. N.; MALAVOLTA, A. T.; MODUGNO, R. G.; VALES, L. F.; CARVALHO, E. G.; EVANGELISTA, S. H.; STEFANI, M. A.; CASTRO NETO, J. C. Focus adjustment method for CBERS 3&4 Satellites MUX camera to be performed in air condition and its experimental verification for best performance in orbital vacuum condition In: INTERNATIONAL CONFERENCE ON SPACE OPTICS, ICSO, 2012, Ajaccio, França. **Proceedings**...Ajjacio: ICSO, 2012.

APÊNDICE A – Revisão da teoria de aberrações de onda aplicáveis a sistemas plano-simétricos

Esta revisão se baseia no trabalho de Sasián (3), que trata da teoria de aberrações de onda aplicável a sistemas plano-simétricos compostos por superfícies esféricas ou ligeiramente asféricas.

Esta teoria será descrita detalhadamente por ser utilizada nas análises apresentadas na seção 3.

Esta seção tem o objetivo de complementar o texto apresentado na seção 2.2.3, que discute os conceitos e resultados obtidos pelo autor no trabalho original. São apresentadas as definições adotadas pelo autor e uma descrição detalhada dos coeficientes das aberrações de onda.

Definições:

- *Optical axis ray,* ou raio do eixo óptico (OAR): é um raio de referência que pertence ao plano de simetria do sistema e exerce o mesmo papel num sistema plano-simétrico que o eixo óptico exerce para o caso de um sistema axialmente simétrico. É o raio que define o centro do campo de visada, o centro do *STOP* de abertura e o centro das pupilas de entrada e saída.
- *Normalized aperture vector,* ou vetor de abertura normalizado ($\boldsymbol{\rho}$): especifica qualquer ponto na abertura do sistema. Possui sua origem no ponto de intersecção do OAR e o plano da pupila de saída.
- *Normalized field vector,* ou vetor de campo normalizado ($\mathbf{H}$): especifica qualquer ponto no campo de visada do sistema. Possui sua origem no ponto de intersecção do OAR e o plano imagem.
- *Unit vector,* ou vetor unitário (i): identifica a direção do plano de simetria (vetor normal ao plano).
- Ângulo I: é o ângulo formado pelo vetor OAR e o vetor normal a uma determinada superfície do sistema.

- Ângulo α: é o ângulo entre o vetor unitário (i) e o vetor de campo ($\mathbf{H}$).
- Ângulo β: é o ângulo entre o vetor unitário (i) e o vetor de abertura ($\boldsymbol{\rho}$).
- Ângulo φ: é o ângulo dado por: $\varphi = \alpha - \beta$, é o ângulo entre o vetor de campo ($\mathbf{H}$) e o vetor de abertura ($\boldsymbol{\rho}$).
- *STOP* de abertura: é a abertura que limita a entrada de luz no sistema. Considerada circular e perpendicular ao OAR.
- Pupila de entrada: é a imagem do *STOP* de abertura no espaço objeto.
- Pupila de saída: é a imagem do *STOP* de abertura no espaço imagem.
- *Chief ray,* ou raio principal: raio do extremo do campo de visada que passa pelo centro do *STOP* de abertura.
- *Marginal ray,* ou raio marginal: raios do centro do campo que passam pela borda do *STOP* de abertura.
- Plano tangencial: é o plano de simetria.
- Plano sagital: é o plano perpendicular ao plano de simetria e que contém o OAR. É, na verdade, um conjunto de planos, já que em cada superfície sua orientação muda de acordo com o OAR.

Coeficiente de aberração:

É necessário traçar dois raios paraxiais para que seja possível calcular os coeficientes, sendo um raio marginal e um raio principal. Definem-se as seguintes grandezas:

- Invariante de Lagrange – Ψ: é dado por:

$$\Psi = n(\bar{u}x - u\bar{x}), \tag{106}$$

onde u representa a inclinação do raio marginal e x indica a altura do raio em uma determinada superfície do sistema. Para o raio principal, esses parâmetros são dados por $\bar{u}$ e $\bar{x}$. Essas grandezas são medidas em relação ao OAR.

- A, conhecida como Invariante de Snell (5), representa o índice de refração de um meio que precede uma superfície multiplicado pelo ângulo de incidência do raio marginal e principal (no caso de $\overline{A}$) e é dado por:

$$A = ni = n\left(u + \frac{x\cos(I)}{R}\right), \overline{A} = n\overline{i} = n\left(\overline{u} + \frac{\overline{x}\cos(I)}{R}\right), \tag{107}$$

  onde R é o raio de curvatura da superfície.

- *Abbe diference operator* $\Delta(\ )$: indica a mudança do argumento na refração, ou seja,

$$\Delta\left(u/n\right) = u/n - u'/n'. \tag{108}$$

- H e H': correspondem às alturas do objeto e da imagem, respectivamente, medidos a partir OAR.

- θ e θ': representam os ângulos de inclinação do plano objeto e do plano imagem.

As seguintes grandezas são definidas:

$$S_I = -\frac{1}{8}A^2\Delta(\frac{u}{n})x, \tag{109}$$

$$S_{II} = -\frac{1}{4R}\Psi^2\Delta\left(\frac{\cos(I)}{n}\right), \tag{110}$$

$$J_I = -\frac{1}{2}n^2\text{sen}^2(\text{I})\Delta\left(\frac{u}{n}\right)x, \tag{111}$$

$$J_{II} = -\frac{1}{2}n\text{sen(I)A}\Delta\left(\frac{u}{n}\right)x, \tag{112}$$

$$J_{III} = -n\text{sen(I)}\Psi\Delta\left(\frac{u}{n}\right), \tag{113}$$

$$J_{IV} = -\frac{1}{2}\frac{n\text{sen}(I)}{R}\Psi\Delta\left(\frac{1}{n}\right)x, \tag{114}$$

$$J_V = -\frac{1}{2}n\text{sen(I)}\Psi^2\Delta\left(\frac{1}{n^2}\right)\frac{1}{x}, \tag{115}$$

$$T_I = -\frac{1}{2}\Psi\Delta[u\mathrm{tg}(\theta)], \tag{116}$$

$$T_{II} = \Psi\Delta[uH\mathrm{tg}(\theta)]\frac{1}{x}. \tag{117}$$

Segundo Sasián, para superfícies asféricas, o *sag* (z) de uma superfície com simetria bilateral é dado por:

$$z = z_{esférico} + z_{asférico}. \tag{118}$$

Para uma superfície esférica, o $z_{esférico}$ é o *sag* da uma esfera base de raio R. O $z_{asférico}$ é o *sag* de uma capa asférica dada por:

$$z_{asférico} = \alpha(\mathbf{i}\cdot\boldsymbol{\rho})^2 + \beta(\mathbf{i}\cdot\boldsymbol{\rho})(\boldsymbol{\rho}\cdot\boldsymbol{\rho}) + \gamma(\boldsymbol{\rho}\cdot\boldsymbol{\rho})^2. \tag{119}$$

Os coeficientes α, β e γ especificam a quantidade de asfericidade de cada tipo que compõe a capa. Os três tipos de asfericidades são: paraboloide cilíndrico (*cilindrical paraboloide*), superfície comática (*comatic surface*) e superfície axialmente simétrica de quarta ordem (*fourth-order axially symmetric surface*).

Para o caso em que o *STOP* está localizado sobre a superfície, a contribuição ΔW para o coeficiente de aberração é dada por:

$$\Delta W = \Delta\left[n\cos(I)\right]z_{asferico}. \tag{120}$$

Para casos em que o *STOP* está distante da superfície:

$$z_\alpha = \alpha\Delta[n\cos(I)]x^2, \tag{121}$$

$$z_\beta = \beta\Delta[n\cos(I)]x^3, \tag{122}$$

$$z_\gamma = \gamma\Delta[n\cos(I)]x^4. \tag{123}$$

A Tabela 30 exibe os coeficientes da equação de aberração de frente de onda. Nesta tabela, são apresentadas as contribuições de superfícies esféricas e asféricas. As três últimas linhas exibem as contribuições devido à inclinação do plano objeto e do plano imagem.

Tabela 30 – Coeficientes das aberrações de frente de onda, para elementos inclinados, contribuição de uma superfície esférica e da capa asférica (3).

Aberração	**Contribuição esférica**		**Contribuição asférica**	
	Coeficiente	Valor	Coeficiente	Valor
Astigmatismo constante (*Constant astigmatism*)	W_{02002}	$\sum_{i=1}^{i=j}\{J_I\}_i$	ΔW_{02002}	z_α
Anamosfismo (*Anamorphism*)	W_{11011}	$\sum_{i=1}^{i=j}\left\{2\frac{\bar{x}}{x}J_I\right\}_i$	ΔW_{11011}	$2\left(\frac{\bar{x}}{x}\right)z_\alpha$
Piston quadrático (*Quadratic piston*)	W_{20020}	$\sum_{i=1}^{i=j}\left\{\left(\frac{\bar{x}}{x}\right)^2 J_I\right\}_i$	ΔW_{20020}	$\left(\frac{\bar{x}}{x}\right)^2 z_\alpha$
Coma constante (*Constant coma*)	W_{03001}	$\sum_{i=1}^{i=j}\{J_{II}\}_i$	ΔW_{03001}	z_β
Astigmatismo linear (*Linear astigmatism*)	W_{12101}	$\sum_{i=1}^{i=j}\left\{2\frac{\bar{x}}{x}J_{II}+J_{III}\right\}_i$	ΔW_{12101}	$2\left(\frac{\bar{x}}{x}\right)z_\beta$
Distorção quadrática I (*Quadratic distortion I*)	W_{21001}	$\sum_{i=1}^{i=j}\left\{\left(\frac{\bar{x}}{x}\right)^2 J_{II}+\frac{\bar{x}}{x}J_{III}+J_V\right\}_i$	ΔW_{21001}	$\left(\frac{\bar{x}}{x}\right)^2 z_\beta$
Inclinação de campo (*Field tilt*)	W_{12010}	$\sum_{i=1}^{i=j}\left\{\frac{\bar{x}}{x}J_{II}+J_{IV}\right\}_i$	ΔW_{12010}	$\left(\frac{\bar{x}}{x}\right)z_\beta$
Distorção quadrática II (*Quadratic distortion II*)	W_{21110}	$\sum_{i=1}^{i=j}\left\{2\left(\frac{\bar{x}}{x}\right)^2 J_{II}+\frac{\bar{x}}{x}(J_{III}+2J_V)\right\}_i$	ΔW_{21110}	$2\left(\frac{\bar{x}}{x}\right)^2 z_\beta$
Piston cúbico (*Cubic piston*)	W_{30010}	$\sum_{i=1}^{i=j}\left\{\left(\frac{\bar{x}}{x}\right)^3 J_{II}+\left(\frac{\bar{x}}{x}\right)^2(J_{III}+J_{IV)}+\frac{\bar{x}}{x}J_V)\right\}_i$	ΔW_{30010}	$\left(\frac{\bar{x}}{x}\right)^3 z_\beta$
Aberração esférica (*Spherical aberration*)	W_{04000}	$\sum_{i=1}^{i=j}\{S_I\}_i$	ΔW_{04000}	z_γ

Continua

Continuação

Aberração		**Contribuição esférica**	**Contribuição asférica**	
Coma linear (Linear coma)	W_{13100}	$\sum_{i=1}^{i=j}\left\{4\frac{\bar{A}}{A}S_I\right\}_i$	ΔW_{13100}	$4\left(\frac{\bar{x}}{x}\right)z_\gamma$
Astigmatismo quadrático (*Quadratic astigmatism*)	W_{22200}	$\sum_{i=1}^{i=j}\left\{4\left(\frac{\bar{A}}{A}\right)^2 S_I\right\}_i$	ΔW_{22200}	$4\left(\frac{\bar{x}}{x}\right)^2 z_\gamma$
Curvatura de Campo (*Field curvature*)	W_{22000}	$\sum_{i=1}^{i=j}\left\{2\left(\frac{\bar{A}}{A}\right)^2 S_I + S_{II}\right\}_i$	ΔW_{22000}	$2\left(\frac{\bar{x}}{x}\right)^2 z_\gamma$
Distorção cúbica (*Cubic distortion*)	W_{31100}	$\sum_{i=1}^{i=j}\left\{4\left(\frac{\bar{A}}{A}\right)^3 S_I + 2\left(\frac{\bar{A}}{A}\right)S_{II}\right\}_i$	ΔW_{31100}	$4\left(\frac{\bar{x}}{x}\right)^3 z_\gamma$
Piston Quadrático (*Quadratic piston*)	W_{40000}	$\sum_{i=1}^{i=j}\left\{\left(\frac{\bar{A}}{A}\right)^4 S_I + \left(\frac{\bar{A}}{A}\right)^2 S_{II}\right\}_i$	ΔW_{40000}	$\left(\frac{\bar{x}}{x}\right)^4 z_\gamma$
Inclinação de campo (*Field tilt*)	ΔW_{12010}	$\sum_{i=1}^{i=j}\left\{\bar{T}\right\}_i$		
Distorção quadrática II (*Quadratic distortion II*)	ΔW_{21110}	$\sum_{i=1}^{i=j}\left\{2\frac{\bar{x}}{x}T_I + T_{II}\right\}_i$		
Piston cúbico (*Cubic piston*)	ΔW_{30010}	$\sum_{i=1}^{i=j}\left\{\left(\frac{\bar{x}}{x}\right)^2 T_I + \frac{\bar{x}}{x}T_{II}\right\}_i$		

Fonte: Adaptada de SASIÁN (3)

APÊNDICE B – Algoritmo *backpropagation* – descrição do processo de ajuste dos pesos sinápticos

Para a descrição do processo de ajuste dos pesos sinápticos, serão apresentados separadamente o processo de ajuste da camada de saída e o processo aplicado às demais camadas da rede.

- Ajuste dos pesos sinápticos da camada de saída:

A camada de saída, definida na Figura 18 como sendo a camada de número 3, possui a matriz de pesos dada por $W_{j,i}^{(3)}$. Da mesma forma que na rede Adaline, o gradiente da função erro é usado no processo de ajuste:

$$\nabla E^{(3)} = \frac{\partial E}{\partial W_{ji}^{(3)}} = \frac{\partial E}{\partial Y_j^{(3)}} \cdot \frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} \cdot \frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}}, \tag{124}$$

onde:

$$\frac{\partial I_j^{(3)}}{\partial W_{ji}^{(3)}} = Y_i^{(2)} \quad , \tag{125}$$

$$\frac{\partial Y_j^{(3)}}{\partial I_j^{(3)}} = \varphi'(I_j^{(3)}) \quad , \tag{126}$$

$$\frac{\partial E}{\partial Y_j^{(3)}} = -(d_j - Y_j^{(3)}) \quad . \tag{127}$$

Sendo $\varphi'(.)$ a derivada de primeira ordem da função de ativação. Assim, tem-se:

$$\frac{\partial E}{\partial W_{ji}^{(3)}} = -(d_j - Y_j^{(3)}) \cdot \varphi'(I_j^{(3)}) \cdot Y_i^{(2)}. \tag{128}$$

Pode-se definir:

$$\delta_j^{(3)} = (d_j - Y_j^{(3)}) \cdot \varphi'(I_j^{(3)}) \quad . \tag{129}$$

O ajuste da matriz de pesos é dado por:

$$\Delta W_{ji}^{(3)} = -\eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)}. \tag{130}$$

Assim, pode-se ajustar os pesos da camada de saída levando-se em consideração o desvio entre as saídas da rede e os valores desejados. A notação algorítmica do processo iterativo é dada por:

$$W_{ji}^{(3)} \leftarrow W_{ji}^{(3)} + \eta \cdot \delta_j^{(3)} \cdot Y_i^{(2)} \quad . \tag{131}$$

- Ajuste dos pesos sinápticos das camadas intermediárias:

No caso dos neurônios das camadas intermediárias, estes não têm acesso direto às saídas da rede; dessa forma, o ajuste de seus pesos sinápticos é realizado considerando-se estimativas dos erros de saída produzidos por aqueles neurônios da camada imediatamente posterior, que já foram ajustados, como o nome diz, os ajustes são feitos da última camada para a primeira. Dessa forma, após o ajuste dos pesos dos neurônios da camada de saída (L=3), ajustam-se os pesos da camada anterior (L=2). Assim, os valores desejados para as saídas de tais neurônios estão disponíveis. Assim, os ajustes serão baseados nos valores ponderados dos pesos que já foram ajustados para a camada imediatamente seguinte, no caso a camada de saída. Para as demais camadas, anteriores a esta, o mesmo processo é repetido.

O ajuste dos pesos sinápticos da camada imediatamente anterior à de saída ($W_{ji}^{(2)}$) é dado pelo gradiente do erro:

$$\nabla E^{(2)} = \frac{\partial E}{\partial W_{ji}^{(2)}} = \frac{\partial E}{\partial Y_j^{(2)}} \cdot \frac{\partial Y_j^{(2)}}{\partial I_j^{(2)}} \cdot \frac{\partial I_j^{(2)}}{\partial W_{ji}^{(2)}}, \tag{132}$$

onde:

$$\frac{\partial I_j^{(2)}}{\partial W_{ji}^{(2)}} = Y_i^{(1)} \quad , \tag{133}$$

$$\frac{\partial Y_j^{(2)}}{\partial I_j^{(2)}} = \varphi'(I_j^{(2)}), \tag{134}$$

$$\frac{\partial E}{\partial Y_j^{(2)}} = \sum_{k=1}^{n_3} \frac{\partial E}{\partial I_k^{(3)}} \cdot W_{kj}^{(3)} = \sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)}. \tag{135}$$

Sendo $W_{kj}^{(3)}$ os pesos sinápticos de todos os neurônios da camada de saída conectados ao *j*-ésimo neurônio da camada anterior, que já foram reajustados e que são usados para o reajuste dos pesos da camada anterior, assim:

$$\frac{\partial E}{\partial W_{ji}^{(2)}} = -(\sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)}) \cdot \varphi'(I_j^{(2)}) \cdot Y_i^{(1)} = \delta_j^{(2)} \cdot Y_i^{(1)}, \tag{136}$$

onde define-se:

$$\delta_j^{(2)} = (\sum_{k=1}^{n_3} \delta_k^{(3)} \cdot W_{kj}^{(3)}) \cdot \varphi'(I_j^{(2)}) \ . \tag{137}$$

Tem-se, assim:

$$\Delta W_{ji}^{(2)} = \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)}. \tag{138}$$

A notação algorítmica do processo iterativo é dada por:

$$W_{ji}^{(2)} \leftarrow W_{ji}^{(2)} + \eta \cdot \delta_j^{(2)} \cdot Y_i^{(1)}. \tag{139}$$

- Ajuste dos pesos sinápticos da primeira camada escondida:

Neste caso tem-se:

$$\nabla E^{(1)} = \frac{\partial E}{\partial W_{ji}^{(1)}} = \frac{\partial E}{\partial Y_j^{(1)}} \cdot \frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} \cdot \frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} \quad , \tag{140}$$

onde:

$$\frac{\partial I_j^{(1)}}{\partial W_{ji}^{(1)}} = x_i, \tag{141}$$

$$\frac{\partial Y_j^{(1)}}{\partial I_j^{(1)}} = \varphi'(I_j^{(1)}), \tag{142}$$

$$\frac{\partial E}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}} \cdot \frac{\partial I_k^{(2)}}{\partial Y_j^{(1)}} = \sum_{k=1}^{n_2} \frac{\partial E}{\partial I_k^{(2)}} \cdot W_{kj}^{(2)}. \tag{143}$$

Assim:

$$\frac{\partial E}{\partial Y_j^{(1)}} = -\sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)}. \tag{144}$$

Tem-se:

$$\frac{\partial E}{\partial W_{ji}^{(1)}} = -(\sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)}) \cdot \varphi'(I_j^{(1)}) \cdot x_i \ = \delta_j^{(1)} \cdot x_i \ , \tag{145}$$

onde define-se:

$$\delta_j^1 = (\sum_{k=1}^{n_2} \delta_k^{(2)} \cdot W_{kj}^{(2)}) \cdot \varphi'(I_j^{(1)}) \quad . \tag{146}$$

Tem-se, assim, que o ajuste dos pesos na direção oposta ao gradiente é:

$$\Delta W_{ji}^{(1)} = \eta \cdot \delta_j^{(1)} x_i \quad . \tag{147}$$

A notação algorítmica do processo iterativo é dada por:

$$W_{ji}^{(1)} \leftarrow W_{ji}^{(1)} + \eta \cdot \delta_j^{(1)} \cdot x_i \quad . \tag{148}$$