ANDERSON NISHIHARA

Montagem Assistida por Realidade Aumentada
Assembly Assisted by Augmented Reality (A³R)

São Paulo
2016

ANDERSON NISHIHARA

Montagem Assistida por Realidade Aumentada
Assembly Assisted by Augmented Reality (A³R)

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de mestre em Ciências

São Paulo
2016

ANDERSON NISHIHARA

Montagem Assistida por Realidade Aumentada
Assembly Assisted by Augmented Reality (A³R)

Dissertação apresentada à Escola Politécnica
da Universidade de São Paulo para obtenção
do título de mestre em Ciências

Área de concentração:
Engenharia de Controle e Automação Mecânica

Orientador:
Jun Okamoto Jr.

São Paulo
2016

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, _____ de _____ de _____

Assinatura do autor: _____

Assinatura do orientador: _____

Catálogo-na-publicação

Nishihara, Anderson

Montagem assistida por realidade aumentada -- Assembly assisted by augmented reality (A³R) / A. Nishihara -- versão corr. -- São Paulo, 2016.
97 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos.

1.montagem do produto (otimização) 2.linha de montagem (processos)
3.visão computacional 4.processamento de imagens I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Mecatrônica e de Sistemas Mecânicos II.t.

AGRADECIMENTOS

Ao Prof. Dr. Jun Okamoto Jr., pela orientação durante todo o desenvolvimento.

Ao Prof. Dr. Marcos de Salles Guerra Tsuzuki, pela sugestão de implementação da solução de reconhecimento de objetos utilizado neste trabalho.

Aos meus colegas do Laboratório de Percepção Avançada pelo companheirismo.

Aos professores, amigos e familiares que contribuíram para minha formação e desenvolvimento desse trabalho.

À empresa Eikon Tecnologia Ltda., por ceder um horário de trabalho flexível para que pudesse me dedicar ao desenvolvimento desse trabalho.

À Escola Politécnica da Universidade de São Paulo, pelas oportunidades abertas durante a graduação e pós-graduação.

À Erica Yuriko Kimura, pela compreensão e apoio durante o desenvolvimento desse trabalho.

RESUMO

Processos de montagem em geral necessitam de instruções para serem executados, desde a montagem de simples brinquedos até máquinas complexas. Tradicionalmente, essas instruções vem na forma de manuais em papel ou meio digital. Seja qual for o modo, os manuais de instruções utilizam desenhos, diagramas ou fotos, além de instruções textuais para indicar a sequência de montagem do início até o estado final. Procurando mudar esse paradigma, esse trabalho propõe um sistema para auxílio à montagem que utiliza realidade aumentada para guiar o usuário no processo. Através de processamento de imagens capturadas por uma câmera o sistema reconhece cada peça e por meio de sinais gráficos é indicado ao usuário qual a peça a ser manipulada e onde deve ser posicionada. Em seguida é feita a verificação do posicionamento das peças e o usuário é alertado quando a tarefa de montagem atinge o estado final. Muitos trabalhos na área utilizam algum tipo de dispositivo customizado como "head mounted display" (HMD) e marcadores para auxiliar o rastreamento da câmera e identificação das peças, limitando a popularização dessa tecnologia. Tendo esse último ponto em vista, propõe-se um sistema que não utiliza qualquer dispositivo customizado ou marcadores para rastreamento. Além disso, todos os processos do sistema são executados em software embarcado, não necessitando de comunicação com outros computadores para o processamento de imagens. Como o sistema não faz uso de marcadores para a identificação das peças, inicialmente é proposto a implementação do sistema para guiar o usuário na resolução de um quebra-cabeças plano. O sistema proposto é denominado como MARA (Montagem Assistida por Realidade Aumentada).

Palavras-chave: montagem do produto (otimização). linha de montagem (processos). visão computacional. processamento de imagens.

ABSTRACT

Assembly processes for simple toys or complex machines usually requires instructions to be executed. Traditionally, these instructions are written in the form of paper or digital manuals. These manuals contains descriptive text, photos or diagrams to guide the assembly sequence from the beginning to the final state. To change this paradigm, it is proposed in this work an augmented reality system to guide assembly tasks. The system recognizes each assembly piece through image processing techniques and guides the piece placement with graphic signals. Later, the system checks if the pieces are properly assembled and warns the user when the assembly have been finished. In the field of assembly assisted by augmented reality systems, many works use some kind of customized device, like head mounted displays (HMD). Furthermore, markers have been used to track camera position and identify assembly parts. These two features restrict the spread of the technology, thus in the proposed work customized devices and markers to track and identify parts shall not be used. Besides, all the processing are executed on embedded software without the need of communication with other computers to help image processing. The first implementation of the proposed system assists the user on the assembly of a planar puzzle, as the proposed system do not use markers to recognize assembly pieces. This system is being called A³R (Assembly Assisted by Augmented Reality).

Keywords: product assembly (optimization). assembly line (processes). computer vision. image processing

LISTA DE ILUSTRAÇÕES

Figura 1	Espaço contínuo de virtualidade segundo Milgram e Kishino. Fonte: (MILGRAM; KISHINO, 1994).	19
Figura 2	Protótipo do sistema de manutenção de impressora laser mostrando como remover o alimentador de papel. Extraído de (FEINER; MACINTYRE; SELIGMANN, 1993).	20
Figura 3	Configuração do ambiente no sistema de Westerfield (2012) com o uso de marcadores para rastreamento. O sistema utiliza um HMD em conjunto com um notebook para auxiliar o usuário na montagem de placa-mãe de um PC. Extraído de (WESTERFIELD, 2012). . .	22
Figura 4	Protótipo de Kitagawa et al. (2011) com uso de projetor. Extraído de (KITAGAWA; YAMAMOTO, 2011).	23
Figura 5	Sistema de RA em arquitetura servidor/cliente. No caso, um PC servidor é utilizado para o processamento de imagens e o <i>smartphone</i> é utilizado como <i>display</i> para exibição da tarefa de montagem. Extraído de (HAKKARAINEN; WOODWARD; BILLINGHURST, 2008). . .	24
Figura 6	Exemplo de marcador fiducial	26
Figura 7	Visão do <i>display</i> do protótipo de Alvarez et al. (2011). O protótipo utiliza rastreamento óptico sem marcadores. Extraído de (ALVAREZ; AGUINAGA; BORRO, 2011).	27
Figura 8	Ferramenta tipo bastão para interagir com elementos virtuais sendo utilizada na montagem de computadores. Extraído de (YUAN; ONG; NEE, 2008).	28
Figura 9	Sistema de Murakami et al. (2013) com uso de interface háptica. Extraído de (MURAKAMI et al., 2013).	29
Figura 10	Protótipo do projeto MOON implementado em plataforma <i>tablet</i> PC. Extraído de (SERVÁN et al., 2012).	31
Figura 11	Diagrama resumo dos trabalhos que utilizam realidade aumentada na montagem. O enquadramento do trabalho proposto está em destaque na cor verde.	32
Figura 12	Montagem de uma lapiseira de 4 peças	34
Figura 13	Exemplo de grafo de conexões de montagem de lapiseira	34
Figura 14	Carregamento dos arquivos de descrição de modelos e sequencia de montagem	38

Figura 15	Diagrama de blocos do sistema proposto	39
Figura 16	Diagrama da descrição dos modelos das peças	40
Figura 17	Diagrama da descrição da sequência de montagem	41
Figura 18	modelo de câmera pinhole	43
Figura 19	Homografia entre plano e imagem	43
Figura 20	Esquema de transformação de perspectiva de um ponto em plano e projeção do ponto em planos de imagem de câmera em duas posições	50
Figura 21	Exemplo de polígono e seu respectivo <i>turning function</i>	53
Figura 22	Exemplo de polígono com outro ponto inicial e seu respectivo gráfico de <i>turning function</i>	54
Figura 23	Polígono rotacionado e seu respectivo gráfico de <i>turning function</i> . . .	54
Figura 24	Diferença da área de dois <i>turning functions</i>	55
Figura 25	Área da diferença entre os <i>turning functions</i>	56
Figura 26	Diagrama de blocos da estimação de submontagens	58
Figura 27	Lente grande angular com distância focal de 18 mm utilizada junto com a câmera do iPad.	60
Figura 28	Pentaminós	61
Figura 29	Esquema da configuração da área de trabalho	62
Figura 30	Foto da configuração da área de trabalho	62
Figura 31	Tela de configuração de parâmetros de algoritmos de segmentação e reconhecimento de objetos	64
Figura 32	Tela de configuração de limiares de cor	65
Figura 33	Padrão tamanho 9 x 6 utilizado na calibração	68
Figura 34	Tela de seleção de sequência de montagem carregada no sistema . . .	70
Figura 35	Esquema de arranjo do <i>tablet</i> e sistema de coordenadas da câmera, global e da câmera virtual	71
Figura 36	Caso 1 - Resultado da remoção de distorção das lentes	73
Figura 37	Caso 2 - Resultado da remoção de distorção das lentes	74
Figura 38	Caso 3 - Resultado da remoção de distorção das lentes	74
Figura 39	Caso 4 - Resultado da remoção de distorção das lentes	74
Figura 40	Imagem sem efeito de distorção das lentes. Observa-se diminuição do contraste próximo das bordas	75
Figura 41	Imagens exemplo de segmentação	77
Figura 42	Imagens exemplo da remoção do efeito de perspectiva	78
Figura 43	Imagens exemplo de reconhecimento das peças do pentaminó	79

Figura 44	Caso em que a peça “U” não foi identificada devido à deformação de sua forma	80
Figura 45	Símbolo de ação de inversão da face da peça "F"	80
Figura 46	Símbolo de ação de inversão da face da peça "L"	81
Figura 47	Símbolo de ação de inversão da face da peça "N"	81
Figura 48	Símbolo de ação de inversão da face da peça "P"	82
Figura 49	Símbolo de ação de inversão da face da peça "Y"	82
Figura 50	Símbolo de ação de inversão da face da peça "Z"	83
Figura 51	Imagens do processo de montagem assistida	84
Figura 52	Montagem de uma solução do quebra-cabeças de tamanho 6x10	85
Figura 53	Montagem de uma segunda solução do quebra-cabeças de tamanho 6x10	86
Figura 54	Montagem de um quebra-cabeças de tamanho 5x12	87
Figura 55	Montagem de um quebra-cabeças de tamanho 4x15	88
Figura 56	Montagem de um quebra-cabeças de tamanho 3x20	89

LISTA DE TABELAS

Tabela 1	Tipos de <i>displays</i> encontrados nas publicações analisadas em RA na montagem	25
Tabela 2	Tipos de rastreamento utilizados nas pesquisas analisadas	27
Tabela 3	Tipos de ferramentas de interação	29
Tabela 4	Plataformas utilizadas pelos sistemas de RA na montagem	31
Tabela 5	Representações de uma sequência de montagem da lapiseira	36
Tabela 6	Resultado da segmentação	76
Tabela 7	Resultado do reconhecimento de objetos	78
Tabela 8	Tempo de execução dos processos	90

LISTA DE LISTAGENS

Listagem 1	Exemplo de parte do arquivo JSON com descrição do modelo da peça "V"	66
Listagem 2	Exemplo de parte do arquivo JSON com descrição do modelo da peça "U"	67
Listagem 3	Exemplo de descrição de sequência de montagem formatada em JSON	69

LISTA DE SÍMBOLOS

M	Ponto 3D de um objeto em coordenada global
m	Ponto na coordenada da imagem
C	Centro óptico de uma lente
n	Ponto 2D da imagem em coordenada homogênea
N	Ponto 3D em coordenada homogênea
R	Matriz de rotação
t	Vetor de translação
A	Matriz de parâmetros intrínsecos da câmera
u_0	Projeção do ponto principal no eixo u
v_0	Projeção do ponto principal no eixo v
α	Fator de escala no eixo u da imagem
β	Fator de escala no eixo v da imagem
γ	Inclinação entre os eixos u e v da imagem
P	Matriz de projeção da câmera
r_1	Primeira coluna da matriz de rotação R
r_2	Segunda coluna da matriz de rotação R
r_3	Terceira coluna da matriz de rotação R
H	Matriz de homografia
D	Ponto de um modelo na coordenada global
d	Projeção de um ponto D no plano da imagem
h_1	Primeira coluna da matriz de homografia H
h_2	Segunda coluna da matriz de homografia H
h_3	Terceira coluna da matriz de homografia H
k_1	Primeiro parâmetro de distorção radial das lentes
k_2	Segundo parâmetro de distorção radial das lentes
k_3	Terceiro parâmetro de distorção radial das lentes
p_1	Primeiro parâmetro de distorção tangencial das lentes
p_2	Segundo parâmetro de distorção tangencial das lentes
G	Intensidade do gradiente no algoritmo Canny Edge
φ	Direção do gradiente no algoritmo Canny Edge
$\phi(s)$	Ângulo de virada do <i>turning function</i>
s	Comprimento de arco do polígono
u	Deslocamento no comprimento da curva do polígono
$D(A, B)$	Distância entre <i>turning function</i> dos polígonos A e B

ρ	Ângulo ótimo entre dois <i>turning functions</i>
η_{msp}	Fator de escala entre objeto segmentado e modelo
H_c	Matriz de homografia da câmera
H_{sp}	Matriz de homografia da imagem sem efeito de perspectiva
h	Distância entre base do <i>tablet</i> e centro geométrico da câmera
θ	Ângulo de inclinação do <i>tablet</i>
h_v	Distância entre plano da câmera virtual e o plano dos objetos

SUMÁRIO

1 INTRODUÇÃO	16
2 REALIDADE AUMENTADA	19
2.1 REALIDADE AUMENTADA NA MONTAGEM	20
2.1.1 Display	21
2.1.2 Rastreamento	24
2.1.3 Interação	27
2.1.4 Plataforma	29
2.2 CONCLUSÃO DO CAPÍTULO	32
3 REPRESENTAÇÃO DA SEQUÊNCIA DE MONTAGEM	33
4 MONTAGEM ASSISTIDA POR REALIDADE AUMENTADA	37
4.1 DESCRIÇÃO DOS MODELOS	40
4.2 DESCRIÇÃO DA SEQUÊNCIA DE MONTAGEM	40
4.3 CORREÇÃO DA DISTORÇÃO DA CÂMERA	41
4.3.1 Geometria de câmera	42
4.3.2 Homografia entre plano e imagem	43
4.3.3 Calibração da câmera	44
4.4 SEGMENTAÇÃO DAS PEÇAS	47
4.4.1 Canny Edge	48
4.5 REMOÇÃO DOS EFEITOS DA PERSPECTIVA	50
4.6 RECONHECIMENTO DE OBJETOS	51
4.6.1 <i>Turning Function</i>	52
4.7 ASSISTÊNCIA NA MONTAGEM	55
4.7.1 Sobreposição da representação gráfica do modelo sobre a peça	56
4.7.2 Estimação de submontagens	57
4.7.3 Validação de conexão das peças	59
5 IMPLEMENTAÇÃO DO SISTEMA DE MONTAGEM ASSISTIDA POR REALIDADE AUMENTADA	60
5.1 CONFIGURAÇÃO DO SISTEMA	63
5.1.1 Calibração da câmera	64
5.2 DESCRIÇÃO DE MODELOS E SEQUÊNCIA DE MONTAGEM ATRAVÉS DE ARQUIVOS	65
5.3 REMOÇÃO DOS EFEITOS DA PERSPECTIVA	68

6 RESULTADOS E ANÁLISES DO SISTEMA DE MONTAGEM ASSIS-	
TIDA POR REALIDADE AUMENTADA	73
6.1 CALIBRAÇÃO DA CÂMERA	73
6.2 SEGMENTAÇÃO DAS PEÇAS	75
6.3 REMOÇÃO DO EFEITO DE PERSPECTIVA	76
6.4 RECONHECIMENTO DAS PEÇAS DE MONTAGEM	77
6.5 MONTAGEM ASSISTIDA	79
6.6 TEMPO DE EXECUÇÃO DOS ALGORITMOS	84
7 CONCLUSÃO	91
Referências	93

1 INTRODUÇÃO

Processos de montagem em geral necessitam de instruções para serem executados, desde a montagem de simples brinquedos até máquinas complexas. Tradicionalmente, essas instruções vem na forma de manuais em papel ou meio digital. Seja qual for o modo, os manuais de instruções utilizam desenhos, diagramas ou fotos, além de instruções textuais para indicar a sequência de montagem do início até o estado final. Procurando mudar esse paradigma, esse trabalho propõe um sistema para auxílio à montagem que utiliza realidade aumentada para guiar o usuário no processo.

Para o auxílio numa tarefa de montagem há vários fatores que devem ser levados em consideração tais como a posição e orientação das peças e sequência de montagem. Além disso, cada tipo de tarefa pode requerer um nível diferente de guia visual e háptico (BOUD et al., 1999) que, devido a dificuldade, pode necessitar de algumas atividades cognitivas sofisticadas. Como forma de auxiliar tarefas de montagem, diversos autores têm proposto sistemas que utilizam realidade aumentada (BOUD et al., 1999; SÄÄSKI et al., 2008; YUAN; ONG; NEE, 2008). O potencial de uso dessa tecnologia é muito grande pois possibilita a sobreposição de instruções de montagem sobre uma imagem em tempo real (ZAUNER et al., 2003; YUAN; ONG; NEE, 2008), além de imagens adicionais que auxiliem na montagem (ZAUNER et al., 2003; SÄÄSKI et al., 2008; WESTERFIELD, 2012; YUAN; ONG; NEE, 2008) e outras informações de maneira a instruir ou mostrar dicas visuais de acordo com o estado de montagem corrente e guiar o usuário passo a passo até atingir o objetivo final.

Trabalhos realizados com a realidade aumentada (RA) no auxílio à montagem propõem que instruções sobrepostas às imagens seriam uma alternativa ao tradicional manual em papel que ajudaria a diminuir o tempo de montagem. Alguns trabalhos (TANG et al., 2003; WIEDENMAIER et al., 2003; BAIRD; BARFIELD, 1999) têm como abordagem a comparação da efetividade entre as várias implementações de sistemas de RA, instruções assistidas por computadores e a tradicional instrução escrita em papel. Os resultados mostraram-se favoráveis ao uso de sistemas de RA. Wiedenmaier et al. (2003) fez uma comparação entre manual em papel, montagem assistida por RA e montagem auxiliada por um especialista. Foi utilizado uma abordagem comparando tarefas em diferentes níveis de complexidade. Concluiu-se que, para tarefas simples, manuais em papel são mais eficientes, contudo, para tarefas complexas, apesar de o auxílio de um especialista ter sido mais eficiente, sistemas de auxílio a montagem utilizando RA obtiveram resultados melhores que manuais em papel.

O presente trabalho trata do desenvolvimento de um sistema de Realidade Aumentada para o auxílio na montagem em tempo real. Este sistema serve como um manual de instruções interativo onde modelos virtuais das partes a serem montadas são colocados sobre imagens

das partes reais e, em uma sequência passo a passo o usuário é guiado até a conclusão do processo de acordo com o estado atual de configuração do espaço de trabalho. A princípio é utilizado um *tablet* que pode ser colocado entre o usuário e a sua área de trabalho de modo que o mesmo pode observar suas mãos e as peças a serem montadas no *display*. Na imagem capturada, o sistema reconhece as peças de montagem através de técnicas de processamento de imagens e, com sinais gráficos sobrepostos na imagem indica ao usuário as peças identificadas e seus respectivos locais de encaixe. Em uma etapa posterior, o sistema verifica se as peças de montagem estão posicionadas nos locais corretos e, assim, ao atingir o estado final de montagem, alerta o usuário que a montagem foi finalizada.

A solução proposta possui as seguintes características:

- Não são utilizados marcadores no mundo real: os objetos são detectados através do processamento de imagens e rotulados de acordo com um banco de dados;
- A solução final da montagem é conhecida: o sistema não é destinado para resolver o problema de montagem em si, e sim, guiar o usuário para que alcance o objetivo final;
- O *display* do dispositivo é posicionado entre o usuário e as peças de montagem e estará fixo durante todo o processo;
- A implementação no *tablet* executa um software totalmente embarcado, utilizando apenas recursos do próprio *tablet* de modo a não utilizar conexões de rede sem fio para comunicação com outros computadores para auxiliar o processamento. Todo o processamento é feito pelo *tablet*.

A maioria das abordagens analisadas utilizam um *display* do tipo *head mounted display* (HMD), o que restringe o uso do sistema de RA para auxílio a montagem a grupos com acesso a este tipo de *display*. Além disso, muitas abordagens utilizam marcadores para auxiliar a detecção e identificação dos objetos da área de trabalho, impondo uma preparação do ambiente de trabalho. A ideia do sistema proposto é que não seja necessário equipamento de alto custo e preparo prévio do ambiente. Desta forma, o usuário recebe um software para ser instalado em dispositivo próprio, como um *tablet* ou *Hololens* (CHEN et al., 2015), e somente isso é o suficiente para iniciar e completar o processo de montagem. As informações sobre as peças e processo de montagem podem ser adquiridos através de arquivo ou pela nuvem, na Internet. O sistema proposto não pretende solucionar o problema de montagem ou ensinar ao usuário o processo de montagem, e sim, guiar qualquer usuário no processo. Assim, no lugar de textos, o sistema proposto faz o uso de linguagem visual, como setas e destaques de área para guiar o usuário na tarefa.

Apesar de muitos trabalhos analisados tratarem de problemas em 3D, o presente trabalho trata somente de problemas em planos para que seja possível o uso do sistema com desempenho em tempo real em equipamento de prateleira sem modificações. Como resultado, implementou-se um protótipo de um sistema de montagem assistido por realidade aumentada em *tablet* para guiar o usuário na montagem de um quebra-cabeças plano sem a utilização de marcadores. Embora a implementação em *tablet* esteja limitado a problemas em planos, o sistema está sendo concebido de forma a possibilitar a implementação em outras plataformas, como numa plataforma do tipo *Hololens*.

2 REALIDADE AUMENTADA

No espaço contínuo de virtualidade de Milgram e Kishino (1994), a realidade aumentada (RA) é classificada como uma realidade mista próxima do ambiente real. RA é a tecnologia que permite a integração entre o mundo real e o virtual através da sobreposição de objetos virtuais ou informações sobre o ambiente real, aumentando o ambiente real. A medida que tecnologias de RA estão amadurecendo, aplicações em diversas áreas, como na educação, medicina, entretenimento, design e engenharia (AZUMA, 1997; AZUMA et al., 2001; KREVELEN; POELMAN, 2010; CARMIGNIANI et al., 2010; CHI; KANG; WANG, 2013) estão surgindo.

O conceito de contínuo de virtualidade (*virtuality continuum*) descrito na figura 1 relaciona várias classes de objetos presentes em um *display*, onde ambientes reais estão em um extremo do contínuo e ambientes virtuais estão no extremo oposto. Os ambientes reais consistem somente de objetos reais como por exemplo o que é observado diretamente do mundo real através de uma câmera. Em ambientes virtuais estão objetos puramente virtuais, como por exemplo uma simulação gráfica de computador. Entre os dois extremos temos a realidade mista, onde objetos reais e virtuais são apresentados juntos em um único *display*.

Figura 1 - Espaço contínuo de virtualidade segundo Milgram e Kishino. Fonte: (MILGRAM; KISHINO, 1994).

Em um sistema RA, o usuário interage com o sistema ao observar objetos virtuais no mundo real através de algum *display*, que pode ser na forma de *head mounted display* (HMD) ou LCD. Existem diversos trabalhos na área, como a aplicação na manutenção de impressoras laser (FEINER; MACINTYRE; SELIGMANN, 1993) e o trabalho de um grupo da Boeing onde a tecnologia de RA estaria sendo implantada para ajudar os técnicos na construção da armadura de cabos elétricos que fazem parte do sistema elétrico de uma aeronave (CAUDELL; MIZELL, 1992).

2.1 REALIDADE AUMENTADA NA MONTAGEM

Há grande interesse em aplicar a realidade aumentada em tarefas de manutenção e montagem. Este interesse é refletido na formação de vários consórcios de pesquisa dedicado ao tópico, como o ARVIKA (FRIEDRICH, 2002) e *Services and Training through Augmented Reality* (RACZYNSKI; GUSSMANN, 2004).

Os primeiros esforços na área de RA na montagem foram o desenvolvimento de um sistema para auxiliar a montagem de um conjunto de cabos de aeronaves (CAUDELL; MIZELL, 1992). As informações derivadas de design de partes e processos foram armazenados em sistemas CAD e estas informações chegam à fábrica em forma de guias, modelos, desenhos, lista de cabeamento e marcação de locais. Há custos e atrasos devido as mudanças de design nas guias, modelos, entre outros elementos e, para resolver problemas com a constante mudança de informações foi construído um sistema de RA utilizando computadores vestíveis e HMD. Os maiores desafios encontrados foram o alinhamento de objetos reais e virtuais e a capacidade computacional dos computadores portáteis disponíveis na época. Mais tarde, outros pesquisadores desenvolveram aplicações de RA na manutenção de impressoras laser (FEINER; MACINTYRE; SELIGMANN, 1993), na montagem do fecho em uma porta de automóvel (REINERS et al., 1998), dentre outros. Na maioria dos casos, os sistemas de RA na montagem são construídos em torno de computadores *desktop* e necessitam que os usuários utilizem algum sistema HMD.

Figura 2 - Protótipo do sistema de manutenção de impressora laser mostrando como remover o alimentador de papel. Extraído de (FEINER; MACINTYRE; SELIGMANN, 1993).

A abordagem tradicional na manufatura é a utilização de desenhos de montagem (*blueprints*) e possivelmente manuais de instruções. Estudos indicam que tarefas de montagem repetitivas não se beneficiam de sistemas de realidade aumentada para auxiliar na montagem (WIEDEN-

MAIER et al., 2003). Apesar disso, a realidade aumentada pode ser uma tecnologia adequada para ambientes de produção de manufatura customizada. Para a utilização da RA neste ambiente, é importante a integração de sistemas de design com as informações de montagem e, para isto, é necessário uma ferramenta de autoria de conteúdo para sistemas RA. Sääski et al. (2008) apresentam uma ferramenta para a autoria de conteúdo para RA, especificamente integrando ferramentas de software de projeto de montagem (*Design for Assembly*) aos sistemas de design (CAD/PDM/PLM). O objetivo principal de seu trabalho foi desenvolver um processo de criação de conteúdo, partindo de sistemas de design p/ instruções de montagem. Em um caso inicial foi apresentada a montagem de um quebra-cabeças 3D utilizando o sistema desenvolvido anteriormente (SALONEN et al., 2007), onde em cada passo o sistema apresentava para o usuário que peça selecionar e utilizava animação para mostrar o lugar a ser colocado.

Existem estudos que avaliam se a realidade aumentada aplicada para auxiliar em tarefas de montagem melhoram o desempenho de seus usuários em relação ao manual em papel, sistemas assistidos por computador, entre outros sistemas. As conclusões suportam a ideia de que a realidade aumentada melhora o desempenho do usuário tanto em relação a instruções em papel (NAKANISHI et al., 2007) como em sistemas assistidos por computador (BAIRD; BARFIELD, 1999) (TANG et al., 2003) e sistemas com animações em realidade virtual (BOUD et al., 1999).

Especificamente em aplicações de RA na montagem, podemos analisar as pesquisas da área sobre 4 aspectos que permitem a implementação do sistema de RA: *display* utilizado, tipo de sistema de rastreamento, modos de interação com usuário e plataforma no qual o sistema é implementado.

2.1.1 Display

Diferentes tipos de *displays* podem ser utilizados em trabalhos de realidade aumentada. Há basicamente três tipos de *displays*: *optical see-through*, onde a camada virtual é mostrada através de lentes e espelhos semitransparentes; *video see-through*, onde o mundo real é mostrado através de uma alimentação de vídeo; e projetor, onde o mundo virtual é projetado diretamente sobre os elementos reais. *Displays optical see-through* utilizam elementos ópticos semi espelhados para dividir os raios de luz, mantendo a resolução do mundo real intacta. Uma outra característica dessa tecnologia é a ausência de efeitos de paralaxe, o que possibilita uma percepção mais natural do ambiente. Contudo, a combinação de elementos virtuais através de espelhos semitransparentes reduz o brilho e contraste da imagem virtual e real. Por outro lado, *displays video see-through* são opções de menor custo e mais fáceis de

serem implementadas. Além disso, como o mundo real já está digitalizado, torna-se mais fácil a adição e remoção de elementos, o que possibilita um controle maior do ambiente. Entretanto, como as câmeras podem ter ângulos diferentes da visão dos olhos do usuário, *displays video see-through* sofrem do efeito paralaxe, além de diminuir a percepção de profundidade. Quanto ao posicionamento, há *displays* que são colocados na cabeça e posicionados próximos aos olhos (HMD - *head mounted displays*), os espaciais, como é o caso dos projetores e os que são posicionados na distância de alcance das mãos (*handheld*).

A maioria dos *displays* utilizados na área de realidade aumentada em processos de montagem são *displays* que precisam serem conectados em computadores do tipo PC, como um monitor, projetor ou HMD. Dentre as publicações analisadas, a maioria utiliza *displays* do tipo HMD conectados a um computador PC.

No sistema *AR Replay* apresentado por Li et al. (2013), para apresentar um vídeo 3D de um tutor para guiar o usuário é utilizado um monitor de PC, contudo, sugere-se o uso de computadores do tipo *tablet PC*. Do mesmo modo, Hou et al. (2013) utilizaram um monitor para guiar o usuário na montagem de tubulações. Apesar do uso do monitor ou projetor para apresentar o vídeo da área de trabalho com RA serem uma alternativa simples e de fácil configuração, não há imersão do usuário na tarefa, necessitando que o usuário troque constantemente de contexto ao olhar para as instruções do *display* e manipular as peças do mundo real.

Figura 3 - Configuração do ambiente no sistema de Westerfield (2012) com o uso de marcadores para rastreamento. O sistema utiliza um HMD em conjunto com um notebook para auxiliar o usuário na montagem de placa-mãe de um PC. Extraído de (WESTERFIELD, 2012).

Tanto os *displays* HMD *optical see-through* como *video see-through* são amplamente utilizados em pesquisas relacionadas com RA em montagem. Como o HMD é um tipo de dispositivo vestível de imersão, o usuário pode utilizar ambas as mãos durante a montagem. Contudo,

grande parte dos HMD disponíveis no mercado são *displays* volumosos, caros e necessitam uma conexão a um computador do tipo PC, restringindo a mobilidade. Mesmo com o uso de PC tipo notebook, há limitação de bateria do dispositivo.

Os *displays* espaciais utilizam projetores, elementos ópticos e outras tecnologias para apresentar informações diretamente sobre os objetos físicos, sem que o usuário necessite carregar um *display*. É a tecnologia mais integrada ao ambiente e a interação pode ser feita simultaneamente por diversas pessoas. Entre as pesquisas que utilizam projetores (CHIANG et al., 2011; KITAGAWA; YAMAMOTO, 2011; FIORENTINO et al., 2014), apenas Kitagawa et al. (2011) utilizam um projetor como um *display* espacial para projetar informações diretamente sobre as peças de um quebra-cabeças 3D proposto pelo *3DUI contest* de 2011. Tanto Chiang et al. (2011) quanto Fiorentino et al. (2014) utilizam projetores apenas como um meio de mostrar o vídeo da área de trabalho aumentado por instruções de montagem para treinamento. Apesar dos projetores serem alternativas para dar a sensação de imersão do usuário na tarefa, dependem de uma iluminação ambiente controlada e necessitam de preparação da área de trabalho para ajustar todos os parâmetros de alinhamento sobre as peças.

(a) Configuração da área de trabalho.

(b) Foto da superfície da área de trabalho. As instruções para solução do puzzle são projetadas diretamente sobre a área de trabalho.

Figura 4 - Protótipo de Kitagawa et al. (2011) com uso de projetor. Extraído de (KITAGAWA; YAMAMOTO, 2011).

Os *displays handheld* são compostos por uma variedade de dispositivos, como *smartphones*, *PDA*, *tablet*, dentre outros dispositivos móveis. Hakkarainen et al. (2008) apresentam como alternativa móvel um sistema que utiliza o *display* de um *smartphone* para guiar o usuário na montagem de um quebra-cabeças 3D, contudo, apesar de ser um dispositivo móvel, ainda depende de uma conexão com um computador PC, pois trabalha dentro de uma arquitetura cliente-servidor.

Figura 5 - Sistema de RA em arquitetura servidor/cliente. No caso, um PC servidor é utilizado para o processamento de imagens e o *smartphone* é utilizado como *display* para exibição da tarefa de montagem. Extraído de (HAKKARAINEN; WOODWARD; BILLINGHURST, 2008).

A tabela 1 resume os diferentes tipos de *displays* analisados nas publicações.

2.1.2 Rastreamento

Um importante conceito em realidade aumentada é a sensibilidade ao contexto (*context awareness*). Quando a realidade aumentada é utilizada apenas para mostrar informações para o usuário sem que necessite alinhar as coordenadas do mundo virtual com o real, esta é chamada de realidade aumentada sem contexto. Os primeiros trabalhos na área utilizam a realidade aumentada sem contexto, ou seja, não há registro de dados e nem estimação da posição das câmeras. Tanto na pesquisa de Boud et al. (1999) quanto nos estudos feitos por Baird et al. (1999), Nakanishi et al. (2007) e Kitagawa et al. (2011) foram utilizados sistemas de realidade aumentada sem contexto. Neste tipo de realidade aumentada são apresentadas apenas imagens estáticas ou animações sem localizações demonstrando ao usuário como deve ser a montagem, semelhantemente a um manual eletrônico, no qual é mostrado na tela o próximo passo da montagem. Neste aspecto, a diferença com um manual eletrônico comum é o acesso mais rápido à informação relevante.

Tabela 1 - Tipos de *displays* encontrados nas publicações analisadas em RA na montagem

TIPO DE DISPLAY	PESQUISAS
Monitor de PC	(TANG et al., 2003)
	(LI; KAMEDA; OHTA, 2013)
	(HOU; WANG; TRUIJENS, 2013)
HMD	(BAIRD; BARFIELD, 1999)
	(BOUD et al., 1999)
	(TANG et al., 2003)
	(ZAUNER et al., 2003)
	(SALONEN et al., 2007; SÄÄSKI et al., 2008)
	(YUAN; ONG; NEE, 2008)
	(WANG et al., 2009; WANG; ONG; NEE, 2013)
	(HENDERSON; FEINER, 2011)
	(WESTERFIELD, 2012)
	(MURAKAMI et al., 2013)
Handheld	(HAKKARAINEN; WOODWARD; BILLINGHURST, 2008)
Projektor	(CHIANG et al., 2011)
	(KITAGAWA; YAMAMOTO, 2011)
	(FIORENTINO et al., 2014)

O rastreamento da câmera e/ou objetos é necessário para a utilização da realidade aumentada com contexto, ou seja, é necessário para sobrepor informações virtuais sobre o mundo real. Um dos desafios do rastreamento é a perda de informação de profundidade ao capturar a imagem através de uma câmera. Alguns trabalhos utilizam sensores adicionais ou múltiplas câmeras para obter informação de profundidade. No estudo de avaliação de *displays* de realidade aumentada apresentado por Tang et al. (2003) foram utilizados sensores do tipo magnético para recuperar a informação de profundidade. Li et al. (2013) utilizaram câmera do tipo RGB-D composta por sensor de profundidade que utiliza laser infravermelho e câmera RGB, enquanto Murakami et al. (2013) utilizam múltiplas câmeras para relacionar características da cena com o uso do algoritmo PTAMM (*Parallel Tracking and Multiple Mapping*), tornando possível a recuperação de informação 3D. Devido ao uso de sensores adicionais, Tang et al. (2003), Li et al. (2013) e Murakami et al. (2013) não precisam lidar com a perda de informação 3D para alinhar o virtual ao mundo real.

Especificamente na área de realidade aumentada na montagem, a abordagem predominante é a utilização de técnicas de visão computacional. Tipicamente, uma imagem 2D é utilizada

em conjunto a um dispositivo de *display* que mostra em tempo real a visão da câmera através da tela. As pesquisas que utilizam câmeras convencionais 2D seguem duas abordagens: rastreamento com o uso de algum tipo de marcador e sem marcador. Henderson e Feiner (2011) utilizam marcadores reflexivos a infravermelho e LED infravermelho para auxiliar o rastreamento. Entre as pesquisas analisadas, grande parcela utilizam marcadores do tipo fiduciais para o rastreamento (ZAUNER et al., 2003; SALONEN et al., 2007; SÄÄSKI et al., 2008; YUAN; ONG; NEE, 2008; HAKKARAINEN; WOODWARD; BILLINGHURST, 2008; WANG et al., 2009; WANG; ONG; NEE, 2013; CHIANG et al., 2011; WESTERFIELD, 2012; NOVAK-MARCINCIN et al., 2012; SERVÁN et al., 2012; HOU; WANG; TRUIJENS, 2013; FIORENTINO et al., 2014). Marcadores fiduciais são figuras padronizadas de alto contraste e formato simples, o que torna simples sua identificação em um ambiente de trabalho. Dessa forma, aplicações que utilizam marcadores fiduciais podem estimar de forma mais simples a posição e orientação da câmera e dos objetos. Entretanto, o uso de marcadores necessita uma preparação do ambiente de trabalho. A figura 6 apresenta um exemplo de marcador fiducial.

Figura 6 - Exemplo de marcador fiducial

O estado da arte do rastreamento óptico com uma única câmera sem uso de marcadores pode ser dividido em dois métodos (BARANDIARAN; BORRO, 2007): rastreamento quadro a quadro e rastreamento por detecção. No rastreamento quadro a quadro a posição atual é calculada utilizando dados do quadro anterior, enquanto no rastreamento por detecção, as características de objetos na imagem são extraídas para realizar a correspondência com características conhecidas de um modelo. É estimada a posição do objeto uma vez que as correspondências são realizadas. As características do modelo podem serem extraídas de imagens ou modelos 3D e podem serem arestas, vértices, planos, dentre outras características. Alvarez et al. (2011) apresentam um sistema com uso de apenas uma câmera para estimar a posição da câmera através da composição de algoritmos de detecção e rastreamento quadro a quadro para reconhecimento de objeto baseado em modelos.

Figura 7 - Visão do *display* do protótipo de Alvarez et al. (2011). O protótipo utiliza rastreamento óptico sem marcadores. Extraído de (ALVAREZ; AGUINAGA; BORRO, 2011).

Tabela 2 - Tipos de rastreamento utilizados nas pesquisas analisadas

			PESQUISA
SISTEMAS DE RASTREAMENTO	Óptico com uma câmera	Com uso de marcador	(ZAUNER et al., 2003)
			(SALONEN et al., 2007; SÄÄSKI et al., 2008)
			(YUAN; ONG; NEE, 2008)
			(HAKKARAINEN; WOODWARD; BILLINGHURST, 2008)
			(WANG et al., 2009; WANG; ONG; NEE, 2013)
			(CHIANG et al., 2011)
			(HENDERSON; FEINER, 2011)
			(WESTERFIELD, 2012)
			(NOVAK-MARCINCIN et al., 2012)
			(SERVÁN et al., 2012)
			(HOU; WANG; TRUIJENS, 2013)
			(FIORENTINO et al., 2014)
		Sem marcador	(ALVAREZ; AGUINAGA; BORRO, 2011)
	Outras abordagens	Sensor magnético	(TANG et al., 2003)
		Câmera RGB-D	(LI; KAMEDA; OHTA, 2013)
		Múltiplas câmeras	(MURAKAMI et al., 2013)

2.1.3 Interação

Além de registrar dados virtuais no mundo real, os sistemas precisam fornecer algum tipo de interface com o mundo real e virtual. Alguns trabalhos utilizam o teclado e mouse de PC (ZAUNER et al., 2003; SALONEN et al., 2007; SÄÄSKI et al., 2008; FIORENTINO et al., 2014) ou teclado de *smartphone* (HAKKARAINEN; WOODWARD; BILLINGHURST, 2008) para interação enquanto outras pesquisas introduzem propostas de interfaces utilizando painéis virtuais selecionáveis por ferramenta 3D do tipo bastão (YUAN; ONG; NEE, 2008), técnicas para reconhecimento de mãos e gestos (WANG et al., 2009; ONG; WANG, 2011) e a utilização de interfaces hápticas (MURAKAMI et al., 2013).

As ferramentas utilizadas para interação em ambientes de realidade aumentada podem

ser de dois tipos: as que utilizam técnicas de visão computacional para a interação e as ferramentas hápticas. Chiang et al. (2011) utilizam marcadores semelhantes aos utilizados para rastreamento para interagir com os elementos virtuais, enquanto Yuan et al. (2008) utiliza técnicas de segmentação para detectar uma ferramenta do tipo bastão para interagir com painéis virtuais. Dando continuidade a pesquisa da ferramenta de Yuan et al. (2008), Wang et al. (2009) e Ong et al. (2011) aperfeiçoaram a técnica para reconhecer mãos e, posteriormente detectar as pontas dos dedos dos usuários para o reconhecimento de gestos para manipular objetos virtuais, contudo, a pesquisa restringiu-se na manipulação e montagem de objetos virtuais em espaço 2D.

Figura 8 - Ferramenta tipo bastão para interagir com elementos virtuais sendo utilizada na montagem de computadores. Extraído de (YUAN; ONG; NEE, 2008).

Tang et al. (2003) mostraram o uso de comando por voz através da interpretação por agente humano. No sistema utilizado por Salonen et al. (2007) e Sääski et al. (2008), além da utilização do teclado do computador são utilizados controles por gestos e voz. O controle por gestos é realizado através da detecção por cor de pele, necessitando um passo anterior de calibração da cor de pele. Através da mão o usuário pode selecionar um menu virtual ao permanecer com a mão em áreas de ativação nos extremos das imagens. O controle por voz é feito através de comunicação TCP/IP com software separado de reconhecimento de comandos simples de palavras em inglês como *next*, *previous*, *next phase* e *previous phase*.

As interfaces hápticas podem ser divididas em cinéticas e táteis. As interfaces cinéticas oferecem forças de resistência, enquanto as interfaces táteis oferecem realimentação na forma de rigidez e temperatura. Luvas são muito utilizadas em pesquisas em realidade virtual, contudo, em casos de realidade aumentada podem impedir o uso das mãos em atividades do mundo real. Murakami et al. (2013) utilizam um sistema de realimentação háptica desenvolvido anteriormente, o qual o usuário precisa vestir para interagir com objetos virtuais e permitir

a atividade em grandes áreas de trabalho.

Figura 9 - Sistema de Murakami et al. (2013) com uso de interface háptica. Extraído de (MURAKAMI et al., 2013).

Tabela 3 - Tipos de ferramentas de interação

FERRAMENTAS DE INTERAÇÃO	PESQUISA
Teclado e/ou mouse	(ZAUNER et al., 2003)
	(SALONEN et al., 2007; SÄÄSKI et al., 2008)
	(FIORENTINO et al., 2014)
Teclado de smartphone	(HAKKARAINEN; WOODWARD; BILLINGHURST, 2008)
Ferramentas com uso de visão computacional	(SALONEN et al., 2007; SÄÄSKI et al., 2008)
	(YUAN; ONG; NEE, 2008)
	(WANG et al., 2009; ONG; WANG, 2011)
	(CHIANG et al., 2011)
Comando de voz	(TANG et al., 2003)
	(SALONEN et al., 2007; SÄÄSKI et al., 2008)
Dispositivo háptico	(MURAKAMI et al., 2013)

2.1.4 Plataforma

Dentre as pesquisas analisadas na área de RA em tarefas de montagem, o computador do tipo PC é a plataforma mais utilizada (BAIRD; BARFIELD, 1999; BOUD et al., 1999; TANG et al., 2003; ZAUNER et al., 2003; SALONEN et al., 2007; SÄÄSKI et al., 2008; YUAN; ONG; NEE, 2008; WANG et al., 2009; WANG; ONG; NEE, 2013; ALVAREZ; AGUINAGA; BORRO, 2011; CHIANG et al., 2011; HENDERSON; FEINER, 2011; KITAGAWA; YAMAMOTO, 2011; WESTERFIELD, 2012; NOVAK-MARCINCIN et al., 2012; MURAKAMI et al.,

2013; LI; KAMEDA; OHTA, 2013; HOU; WANG; TRUIJENS, 2013; FIORENTINO et al., 2014). A versatilidade de computadores PC, a diversidade de conexões que podem serem feitas com vários dispositivos simultaneamente, a ampla disponibilidade no mercado e o poder de processamento de dados explicam a preferência por este tipo de plataforma. Além disso, há a restrição imposta pelos *displays* utilizados (projetores e HMD's) nos sistemas de RA, pois estes necessitam estarem conectados a computadores.

Um dos requisitos básicos de sistemas de RA é que o tempo de resposta seja mais próximo possível do real. Devido a manipulação de elementos virtuais 3D, os sistemas de RA requerem grande poder de processamento, que pode ser encontrado em plataformas do tipo computadores PC. Além disso, sistemas de rastreamento sem uso de marcadores necessitam de processamento e memória adequados, como no caso de Alvarez et al. (2011), onde foi utilizado computador com processador tipo *Intel Core i7* equipado com 3GB de memória RAM. O uso de marcadores para o rastreamento é uma forma de diminuir a complexidade dos algoritmos de rastreamento reduzindo a demanda por alto poder de processamento.

Ainda são poucas as abordagens de sistemas de realidade aumentada para auxílio a montagem utilizando *tablets* e/ou *smartphones* para torná-los portáteis, pois há limitação de poder de processamento destes dispositivos. Na pesquisa de Hakkarainen et al. (2008) é apresentado um sistema de RA baseado em arquitetura do tipo cliente-servidor chamado *Sagasse (Symbian Augmented Assembly)* onde as informações complexas de um modelo localizam-se em um PC (servidor) e o smartphone é utilizado como um *thin-client* para aumentar imagens estáticas de uma área de trabalho com sequência animadas de montagem. Para demonstrar o sistema em funcionamento foi utilizado um quebra-cabeças 3D e para auxiliar a estimação da posição da câmera e o registro dos modelos nas imagens foram utilizados marcadores sobre a área de trabalho. No sistema, o usuário utiliza a câmera do smartphone para enviar uma foto ao servidor através de *Wifi* ou *Bluetooth*. O servidor faz o processamento da imagem, registrando as animações de montagem que são enviadas de volta para o smartphone a fim de serem exibidas. A arquitetura do tipo cliente-servidor foi escolhido devido a pouca capacidade de renderização 3D do smartphone disponível na época (Nokia N95) e, por esta razão, as informações virtuais não são alinhadas ao mundo real em tempo real. No protótipo utilizado no projeto MOON (*assembly Oriented authOring augmeNted reality*) desenvolvido pela *AIRBUS Military* e apresentado por Servan et al. (2012), marcadores são utilizados para rastrear a câmera e elementos CAD são sobrepostos sobre imagens da área de trabalho capturadas em tempo real pela *webcam* de um *tablet* do tipo PC para auxiliar a localização das conexões elétricas de parte de uma aeronave. Pode-se observar pela tabela 4 que, mesmo com abordagens utilizando dispositivos móveis ainda há a dependência de um PC para o processamento. Entretanto, os dispositivos móveis têm apresentado grandes avanços tecnológicos nos últimos anos. Um representante

deste avanço tecnológico é o *iPad*, um dispositivo móvel que destaca-se pela sua capacidade computacional, o que permite executar algoritmos de processamento de imagem, em tempo real, no próprio dispositivo e pode ser utilizado em sistemas de RA (CARMIGNIANI et al., 2010). Deste modo, é esperado que o uso de plataforma móvel na realidade aumentada cresça juntamente com o avanço do poder de processamento dos dispositivos móveis.

Figura 10 - Protótipo do projeto MOON implementado em plataforma *tablet* PC. Extraído de (SERVÁN et al., 2012).

Tabela 4 - Plataformas utilizadas pelos sistemas de RA na montagem

PLATAFORMA	PESQUISA
PC	(BAIRD; BARFIELD, 1999)
	(BOUD et al., 1999)
	(TANG et al., 2003)
	(ZAUNER et al., 2003)
	(SALONEN et al., 2007; SÄÄSKI et al., 2008)
	(YUAN; ONG; NEE, 2008)
	(WANG et al., 2009; WANG; ONG; NEE, 2013)
	(ALVAREZ; AGUINAGA; BORRO, 2011)
	(CHIANG et al., 2011)
	(HENDERSON; FEINER, 2011)
	(KITAGAWA; YAMAMOTO, 2011)
	(WESTERFIELD, 2012)
	(NOVAK-MARCINCIN et al., 2012)
	(MURAKAMI et al., 2013)
	(LI; KAMEDA; OHTA, 2013)
	(HOU; WANG; TRUIJENS, 2013)
	(FIORENTINO et al., 2014)
Smartphone e PC	(HAKKARAINEN; WOODWARD; BILLINGHURST, 2008)
Tablet	(SERVÁN et al., 2012)

2.2 CONCLUSÃO DO CAPÍTULO

Neste capítulo foram analisados trabalhos que utilizam a realidade aumentada nas tarefas de montagem sob 4 diferentes aspectos: tipo de *display* utilizado, tipo de sistema de rastreamento, interação com usuário e plataforma de implementação. A figura 11 apresenta o número de publicações analisadas sob cada um dos aspectos. O *display* utilizado neste trabalho é do tipo *handheld* e não é utilizado um sistema de rastreamento, pois o tablet mantém-se fixo durante todo o processo de montagem. Para a interação, apenas ferramentas com uso de visão computacional são utilizadas, como a exibição de instruções de montagem através do uso do reconhecimento de objetos. Finalmente, como plataforma de implementação foi escolhido o *tablet*. A figura 11 mostra os diversos aspectos analisados e como este trabalho se enquadra dentro desta análise, destacado em cor verde na figura.

Figura 11 - Diagrama resumo dos trabalhos que utilizam realidade aumentada na montagem. O enquadramento do trabalho proposto está em destaque na cor verde.

3 REPRESENTAÇÃO DA SEQUÊNCIA DE MONTAGEM

Representações de sequência de montagem foram descritas em (DE MELLO; SANDERSON, 1991; GOTTIPOLO; GHOSH, 2003). Um meio de descrever uma representação de sequência de montagem é necessário sempre que um sistema de controle inteligente para processos de montagem é desenvolvido. A escolha da representação impacta na memória necessária para armazenar a informação e na dificuldade de derivar a representação a partir de um processo de montagem.

As seguintes definições são utilizadas para a representação de sequência de montagem (DE MELLO; SANDERSON, 1991):

- *Submontagem* – representa um conjunto de peças conectadas, pode conter uma ou mais peças conectadas;
- *Tarefa de montagem* – é a ação de conexão de submontagens;
- *Sequência de montagem* – lista ordenada de tarefas de montagem;
- *Estado de montagem* – lista de submontagens conectadas.

Uma montagem mecânica é a conexão entre peças sólidas formando uma unidade estável. Cada peça é um objeto rígido, sólido e a sua forma mantém-se inalterada. As peças de montagem são conectadas sempre que uma ou mais superfícies compatíveis estão em contato, o que reduz o grau de liberdade relativo entre as peças. Por exemplo, um contato cilíndrico previne qualquer movimento relativo o qual não é uma translação ao longo do eixo ou uma rotação no eixo. Além disso, anexos também podem agir em superfícies de contato e eliminar todos os graus de liberdade para movimentos relativos.

Uma submontagem é um conjunto contendo uma ou mais peças de modo que, todas as peças pertencem ao conjunto estão interconectadas. No início da montagem o número de submontagens é o mesmo que o número de peças e, ao final da montagem, o número de submontagens é igual a 1. Duas submontagens podem ser conectadas formando uma submontagem maior. Quando uma submontagem é formada, todos os contatos de superfície que formam a conexão são estabelecidos, deste modo, uma tarefa de montagem é executada.

Uma lista ordenada de tarefas ou estados de montagem é uma sequência de montagem. Um processo de montagem inicia com todas as peças separadas e termina com todas as peças unidas para formar um objeto.

Uma montagem pode ser representada por um grafo sem orientação (P, C) , denominado grafo de conexões, no qual $P = \{p_1, p_2, p_3, \dots, p_N\}$ é um conjunto de vértices e $C = \{c_1, c_2, c_3, \dots, c_L\}$ é um conjunto de arestas. Cada vértice em P representa uma parte ou

Figura 12 - Montagem de uma lapiseira de 4 peças

peça da montagem e, cada aresta em C representa todos os contatos de superfície entre duas peças para formar uma conexão. C é denominado conexão e representa ao menos uma superfície geométrica de cada peça em contato. A figura 13 apresenta um exemplo de grafo de conexões de montagem da lapiseira gerado a partir da montagem de uma lapiseira ilustrado na figura 12.

Figura 13 - Exemplo de grafo de conexões de montagem de lapiseira

O estado de montagem representa a configuração do processo de montagem no início de uma tarefa de montagem. A configuração das peças é dada pelas conexões que já foram estabelecidas. Deste modo, podemos representar o estado de montagem através de um vetor binário $\bar{x} = [x_1 x_2 \dots x_L]$ em que cada componente x_i representa uma conexão estabelecida. Se a i – *enésima* conexão está estabelecida, então o seu valor é T . O valor é F caso contrário.

Por exemplo, o estado inicial de montagem da figura 12 é $[FFFFF]$, ou seja, no estado inicial não há nenhuma conexão estabelecida. O estado de montagem resultante da conexão c_1 é $[TFFFF]$. Ao estabelecermos a conexão c_2 , a conexão c_3 também deve ser estabelecida, portanto, o vetor binário que representa este estado é $[TTTFF]$. O estado final da montagem é alcançado quando todas as conexões estão estabelecidas $[TTTTT]$.

Como a submontagem foi definida como sendo o conjunto de uma ou mais peças cujas as interconexões estão todas estabelecidas, a submontagem pode ser representada por um subconjunto de peças da montagem. Deste modo, qualquer estado de montagem também

pode ser representado por uma partição do conjunto de peças da montagem. Por exemplo, o estado inicial da montagem da lapiseira pode ser representado por $\{\{Corpo\}, \{Mecanismo\}, \{Cabeça\}, \{Capa\}\}$. Ao conectar o corpo ao mecanismo, o estado resultante é representado por $\{\{Corpo, Mecanismo\}, \{Cabeça\}, \{Capa\}\}$. Ao final do processo, o estado da montagem é $\{\{Corpo, Mecanismo, Cabeça, Capa\}\}$.

Nem todo subconjunto de peças da montagem representa um estado de montagem. A partição $\{\{Cabeça, Capa\}, \{Corpo\}, \{Mecanismo\}\}$ não caracteriza um estado de montagem, pois não há conexão possível entre estas peças. Do mesmo modo, nem todo vetor binário de conexões representa um estado de montagem. $[TTFFF]$ não pode representar um estado de montagem, pois se as conexões c_1 e c_2 estão estabelecidas, a conexão c_3 também tem que estar estabelecida.

Se duas submontagens caracterizadas por seus correspondentes conjuntos de peças θ_i e θ_j forem unidas e o conjunto final de peças θ_k formar uma submontagem possível, a ação realizada é uma tarefa de montagem. Por exemplo, na montagem mostrada na figura 12, se $\theta_i = \{Corpo\}$ e $\theta_j = \{Mecanismo\}$, então a junção de θ_i e θ_j é uma tarefa de montagem. A caracterização desta tarefa de montagem pode ser realizada através das submontagens $\{\{Corpo\}, \{Mecanismo\}\}$. Ainda, uma tarefa de montagem pode ser decomposta em uma submontagem e o conjunto de conexões estabelecidas para gerar a submontagem. Por exemplo, na figura 12, a junção da cabeça à submontagem corpo e mecanismo pode ser representado por $\{\{Cabeça, Corpo, Mecanismo\}, \{c_2, c_3\}\}$.

Se assumirmos que as submontagens já formadas não são desfeitas durante um processo de montagem, uma lista ordenada de tarefas de montagem pode caracterizar um processo de montagem do início ao fim. Assim, uma sequência de montagem pode ser representada de diversos modos:

- Lista ordenada de tarefas de montagem. O número de elementos na lista é igual ao número de peças menos 1;
- Lista ordenada de vetores binários. Cada vetor corresponde a um estado de montagem. O número de elementos nesta lista é igual ao número de peças;
- Lista ordenada de partições do conjunto de peças de montagem. Cada partição corresponde a um estado de montagem. O número de elementos desta lista é igual ao número de peças;
- Lista ordenada de subconjuntos de conexões. O número de elementos desta lista é igual ao número de peças menos 1.

Por exemplo, seja a seguinte sequência de montagem da figura 12 e grafo 13:

1. Conectar o corpo ao mecanismo da lapiseira através da conexão c_1 , resultando na submontagem $\{Corpo, Mecanismo\}$;
2. Estabelecer as conexões c_2 e c_3 , conectando as submontagens $\{Corpo, Mecanismo\}$ e $\{Cabeça\}$. Ao final da tarefa, a submontagem $\{Corpo, Mecanismo, Cabeça\}$ é formada;
3. Estabelecer as conexões c_4 e c_5 , conectando as submontagens $\{Corpo, Mecanismo, Cabeça\}$ e $\{Capa\}$. Ao final da tarefa o processo de montagem é finalizado e há apenas uma submontagem $\{Corpo, Mecanismo, Cabeça, Capa\}$.

Na tabela 5 são representados os 4 tipos de sequência de montagem do exemplo acima.

Tabela 5 - Representações de uma sequência de montagem da lapiseira

Tipo de representação de sequência	Tamanho da lista	Representação
Lista de tarefas de montagem	3	$[\{\{Corpo\}, \{Mecanismo\}\}]$ $\{\{Corpo, Mecanismo\}, \{Cabeça\}\}$ $\{\{Corpo, Mecanismo, Cabeça\}, \{Capa\}\}$
Lista de vetores binários	4	$[FFFFF]$ $[TFFFF]$ $[TTTFF]$ $[TTTTT]$
Lista de partições do conjunto de peças da montagem	4	$[\{\{Corpo\}, \{Mecanismo\}, \{Cabeça\}, \{Capa\}\}]$ $\{\{Corpo, Mecanismo\}, \{Cabeça\}, \{Capa\}\}$ $\{\{Corpo, Mecanismo, Cabeça\}, \{Cabeça\}, \{Capa\}\}$ $\{\{Corpo, Mecanismo, Cabeça, Capa\}\}$
Lista de subconjuntos de conexões	3	$[\{c_1\}, \{c_2, c_3\}, \{c_4, c_5\}]$

Muitos processos de montagem possuem mais de uma sequência de montagem possível e, portanto, vários subconjuntos de elementos da lista de sequência de montagem podem ser repetidos. Na literatura existem representações para descrever as várias sequências de um processo de montagem de forma mais compacta, como grafos direcionados, grafos AND/OR (DE MELLO; SANDERSON, 1991) e tabela de sequência de montagem (GOTTIPOLU; GHOSH, 2003). Estas representações de sequências de montagem são necessárias para que sistemas inteligentes possam verificar todas as montagens possíveis e selecionar a sequência de montagem mais adequada ao contexto de trabalho.

4 MONTAGEM ASSISTIDA POR REALIDADE AUMENTADA

Está sendo proposto nesse trabalho um sistema de realidade aumentada para auxílio em processos de montagem. O sistema pode ser executado através de software embarcado em dispositivos de mercado com câmera, tal como o *iPad*, e sem a necessidade de sua customização, o que permite a popularização da tecnologia até então restrita a ambientes controlados de laboratório e alguns casos de uso em indústrias. Assim, o sistema dispensa o uso de dispositivos customizados, como o HMD das publicações analisadas (tabela 1).

No sistema proposto, o dispositivo de assistência é disposto na área de trabalho entre o usuário e as peças, de forma que a sua câmera capture imagens das peças de montagem em perspectiva. O software do sistema pode ser baixado para o dispositivo de assistência do usuário através de uma loja de aplicativos e todas as informações necessárias para guiar o usuário na montagem podem ser obtidas através da internet, como demonstrado na figura 14. Em processos de montagem as peças são conhecidas, portanto, pode-se utilizar o conhecimento da forma da peça para a identificação das peças na área de trabalho. Deste modo, o sistema reconhece as peças através de algoritmo de reconhecimento de objetos baseado em modelos. O sistema proposto não faz uso de marcadores para o reconhecimento de objetos ou para rastreamento, o que torna o reconhecimento independente da informação da cena e elimina a necessidade de preparação do ambiente para guiar o usuário na montagem.

Durante o processo de montagem, o dispositivo de assistência permanece fixo e o sistema identifica e mostra ao usuário uma representação gráfica do modelo da peça associada à tarefa de montagem corrente no seu *display*. Através de sinais gráficos, como setas e realce de áreas com cores semitransparentes, o sistema indica ao usuário o local de encaixe da peça. A cada tarefa de montagem executada, um sinal sonoro é reproduzido após a verificação do encaixe.

Na fase de inicialização do sistema, os modelos com a descrição da forma das peças e a sequência de montagem são carregados no sistema através da leitura de arquivos de texto. Ambos arquivos podem ser obtidos através da internet, como ilustrado na figura 14. Uma vez com os modelos carregados, é necessário descrever os modelos em um formato invariante a translação, rotação e escala para que seja possível a identificação das peças. Para cada modelo de peça é derivada sua representação invariante a translação, rotação e escala e este é armazenado em banco de dados para posterior utilização no reconhecimento de peças na cena capturada pela câmera do dispositivo. Ainda durante a inicialização é feita a calibração da câmera utilizando o método descrito por (ZHANG, 1999). A calibração da câmera tem como objetivo estimar os parâmetros intrínsecos da câmera e de distorção das lentes. Ambos parâmetros são necessários para relacionar as coordenadas globais no mundo real com as coordenadas das imagens capturadas por uma câmera de modelo *pinhole*.

Figura 14 - Carregamento dos arquivos de descrição de modelos e sequência de montagem

A figura 15a demonstra o algoritmo de processo de montagem. Inicialmente, o sistema obtém a tarefa de montagem a partir da lista de sequência de montagem obtido na inicialização. Cada tarefa de montagem é composta por um identificador de peça da montagem e as suas respectivas conexões. A peça associada com a tarefa é identificada no subsistema de reconhecimento de objetos. Esta identificação é realizada ao comparar a forma da peça com o banco de dados de modelos. Após a identificação da peça, é estimada a posição e orientação da peça e uma representação gráfica do modelo da peça é apresentada na tela do dispositivo de assistência para mostrar ao usuário a peça que deve ser deslocada. Além da representação gráfica do modelo, símbolos gráficos são exibidos para indicar as seguintes ações: inverter a face da peça e conectar a peça à submontagem. O sistema constantemente indica ao usuário a peça e verifica a conexão da peça da tarefa com a submontagem. Quando a conexão de peça com a submontagem é detectada, o sistema verifica se a tarefa de montagem corrente é a última da sequência de montagem. Caso não seja a última tarefa, o processo é repetido com a obtenção da próxima tarefa de montagem até que se obtenha a última tarefa da sequência de montagem. Quando é validada a execução da última tarefa de montagem, o usuário é alertado sobre o fim do processo de montagem.

Para cada quadro de imagem do vídeo, uma sequência de algoritmos de processamento de imagens é executada no subsistema de reconhecimento de objetos, como mostrado na figura 15b. As imagens capturadas pela câmera possuem distorções provocadas pelas lentes. Para que seja possível o reconhecimento das peças, é necessário, primeiramente, remover o efeito de distorção das lentes. O primeiro algoritmo do subsistema utiliza os parâmetros de distorção das lentes, previamente calculados na fase de inicialização, para corrigir as distorções na imagem. Após a remoção da distorção das lentes, os objetos da área de trabalho são segmentados. A segmentação é realizada através de uma série de algoritmos: segmentação por limiar de

cor como filtro inicial, *Canny edge* (CANNY, 1986) para a detecção de bordas das peças, fechamento morfológico para fechar as bordas abertas, *Border following* (SUZUKI; ABE, 1985) para vetorização das bordas em pontos e transformação da lista de pontos em lista de vértices. Devido a posição e orientação da câmera do *tablet*, as imagens capturadas das peças estão em perspectiva. Utilizando os parâmetros intrínsecos da câmera, o efeito da perspectiva é removido das peças segmentadas através da relação das homografias. Em seguida, características invariantes a translação, rotação e escala são calculadas a partir dos contornos das peças segmentadas. A identificação das peças é realizada a partir da comparação das características extraídas das peças segmentadas com as características dos modelos do banco de dados. Uma vez que a peça é identificada, sua localização e orientação são estimados para que seja possível sobrepor a representação gráfica do modelo sobre a peça.

Figura 15 - Diagrama de blocos do sistema proposto

Na seções a seguir serão explicados com detalhes cada processo descrito nos diagramas das

figuras 15a e 15b.

4.1 DESCRIÇÃO DOS MODELOS

Em um processo de montagem as peças são conhecidas previamente. Para que seja possível o reconhecimento de objetos e a sobreposição de representação gráfica da peça, na fase de inicialização do sistema, os modelos das peças são carregados através de leitura de um arquivo. Cada modelo de peça descrito pelo arquivo possui um rótulo e uma lista de arestas geométricas.

O rótulo da peça é um identificador único utilizado para o reconhecimento de objetos e na descrição da sequência de montagem. Na lista de arestas da peça, cada aresta é composto por um número sequencial com início em 0 e um par de vértices geométricos. O número sequencial identifica cada aresta da peça, sendo posteriormente utilizado na descrição da sequência de montagem para identificar as arestas de cada conexão de peças. Cada vértice é composto por um par de coordenadas em uma base arbitrária definida para cada modelo.

A figura 16 apresenta em forma de diagrama a descrição dos modelos das peças de montagem.

Figura 16 - Diagrama da descrição dos modelos das peças

4.2 DESCRIÇÃO DA SEQUÊNCIA DE MONTAGEM

Além do carregamento da descrição dos modelos é necessário o carregamento da descrição da sequência de montagem. Derivou-se uma descrição de sequência de montagem baseado na

tabela de sequência de montagem apresentado por Gottipolu e Ghosh (2003). A sequência de montagem é representada por uma lista ordenada de tarefas de montagem e cada tarefa de montagem possui um número sequencial representando a ordem da tarefa, a informação da peça a ser conectada em uma submontagem, identificada pelo seu rótulo, e a descrição das conexões a serem realizadas na tarefa. Cada conexão é representada por um par de arestas. O par de arestas representa uma aresta da peça associada à tarefa e uma aresta de alguma peça pertencente a submontagem. Desse modo, uma tarefa de montagem é completada toda vez que for detectado que todas as arestas de cada conexão estão concatenadas. Nessa descrição da sequência de montagem, cada tarefa de montagem descreve uma tarefa onde uma peça é manipulada e conectada a uma submontagem formada por peças já manipuladas e conectadas anteriormente.

A figura 17 apresenta em forma de diagrama a descrição da sequência de montagem.

Figura 17 - Diagrama da descrição da sequência de montagem

4.3 CORREÇÃO DA DISTORÇÃO DA CÂMERA

As câmeras disponíveis atualmente possuem lentes que permitem que o sensor receba maiores quantidades de luz, contudo, a utilização de lentes impõem distorções nas imagens obtidas. Para que seja possível o registro de objetos virtuais e reconhecimento de objetos, deve-se corrigir os desvios provocados pelas lentes e conhecer a geometria que permite a transformação de pontos 3D (mundo real) em 2D (imagem). A calibração da câmera é um processo com o qual é possível a estimação da geometria da câmera e os coeficientes da distorção provocada

pelas lentes, ou seja, possibilita relacionar as medidas de unidade das imagens (*pixels*) com as medidas do mundo real. Uma vez estimados os parâmetros da geometria da câmera e os coeficientes de distorção, estes são armazenados para posteriormente serem utilizados na correção de distorções provocadas pelas lentes, remoção do efeito de perspectiva, sobreposição da representação gráfica dos modelos sobre as peças e indicação da posição e orientação da peça na submontagem.

A seguir serão descritas a geometria de uma câmera *pinhole*, homografia e a calibração de câmera para a estimação dos parâmetros característicos da câmera e coeficientes de distorção das lentes.

4.3.1 Geometria de câmera

A câmera *pinhole* modela o processo de projeção de um ponto 3D global M com um ponto da imagem m . A imagem de um ponto M em 3D, indicado por m é formada por um raio óptico que tem origem em M e passa pelo centro óptico C , interceptando o plano da imagem. Os três pontos M , m e C são colineares. Na figura 18, para fins ilustrativos, o plano da imagem é posicionado entre o ponto da cena e o centro óptico, que é matematicamente equivalente a configuração física onde o plano da imagem está localizado antes do centro óptico. A relação entre o ponto M e sua imagem de projeção m é dado por:

$$sn = A \begin{bmatrix} R & t \end{bmatrix} N \equiv PN \quad (1)$$

com

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

e

$$P = A \begin{bmatrix} R & t \end{bmatrix} \quad (3)$$

onde n e N são os pontos m e M , respectivamente, em coordenadas homogêneas; s é um fator de escala arbitrário; $\begin{bmatrix} R & t \end{bmatrix}$ são a rotação e translação que relaciona o sistema de coordenada global com o sistema de coordenada da câmera (também chamados de parâmetros extrínsecos da câmera); A é a matriz intrínseca da câmera, com u_0 e v_0 sendo as coordenadas do ponto principal, α e β são os fatores de escala na imagem de eixos u e v , e γ é o parâmetro que descreve a inclinação entre os dois eixos da imagem. A matriz 3x4 P é chamada de matriz de projeção da câmera. Geralmente o fator γ de inclinação é igual a 0 na maioria das câmeras.

A calibração determina os parâmetros de transformação entre um objeto em 3D e a imagem 2D observada pela câmera através de informação visual. A transformação inclui:

- Parâmetros extrínsecos: orientação (rotação) e localização (translação) da câmera (R, t);
- Parâmetros intrínsecos: características da câmera ($\alpha, \beta, \gamma, u_0, v_0$).

Apesar da matriz de rotação consistir de 9 elementos, esta possui apenas 3 graus de liberdade. O vetor de translação t possui 3 parâmetros. Portanto, há 6 parâmetros extrínsecos e 5 parâmetros intrínsecos, somando um total de 11 parâmetros.

Figura 18 - modelo de câmera pinhole

4.3.2 Homografia entre plano e imagem

Figura 19 - Homografia entre plano e imagem

Sem perda de generalidade, assumimos que um modelo plano está em $Z = 0$ do sistema de coordenada global. Vamos denotar r_1 , r_2 e r_3 como colunas da matriz de rotação R . Da

equação 1:

$$sn = A \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \quad (4)$$

Temos que:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (5)$$

Deste modo, podemos definir a matriz H da homografia como:

$$H = A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (6)$$

Portanto, podemos relacionar um ponto do plano de um modelo com o ponto de sua imagem através da equação abaixo:

$$sd = HD \quad (7)$$

onde d e D são, respectivamente, o ponto da imagem e o ponto do plano do modelo.

4.3.3 Calibração da câmera

Um aspecto importante na realidade aumentada é a sobreposição de informações sobre os objetos reais. Para que isto seja possível, deve-se ter uma relação de pontos 3D com os pontos 2D de uma imagem. Portanto, é necessário o conhecimento dos parâmetros intrínsecos e extrínsecos da câmera. A extração da informação dos parâmetros intrínsecos e extrínsecos é conhecido como calibração da câmera. Há várias técnicas para a calibração da câmera, contudo, iremos focar na técnica de calibração de câmera utilizando imagens de padrões localizados em um plano (ZHANG, 1999).

4.3.3.1 Restrições nos parâmetros intrínsecos

Dada uma imagem de um plano do modelo, uma homografia $H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}$ pode ser estimada. A partir da equação 6, temos:

$$\begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = \lambda A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (8)$$

onde λ é um escalar arbitrário. Sabendo que r_1 e r_2 são ortonormais, temos:

$$h_1^T A^{-T} A^{-1} h_2 = 0 \quad (9)$$

$$h_1^T A^{-T} A^{-1} h_1 = h_2^T A^{-T} A^{-1} h_2 \quad (10)$$

Estas são as duas restrições dos parâmetros intrínsecos, dada uma homografia. Como uma homografia possui 8 graus de liberdade e há 6 parâmetros extrínsecos (3 para rotação e 3 para translação), podemos obter apenas 2 restrições nos parâmetros intrínsecos.

4.3.3.2 Solução fechada da calibração da câmera

Nesta seção demonstraremos como solucionar o problema da calibração da câmera analiticamente. Seja:

$$B = A^{-T} A^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2 \beta} & \frac{v_0 \gamma - u_0 \beta}{\alpha^2 \beta} \\ -\frac{\gamma}{\alpha^2 \beta} & \frac{\gamma^2}{\alpha^2 \beta^2} + \frac{1}{\beta^2} & -\frac{\gamma(v_0 \gamma - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} \\ \frac{v_0 \gamma - u_0 \beta}{\alpha^2 \beta} & -\frac{\gamma(v_0 \gamma - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} & \frac{(v_0 \gamma - u_0 \beta)^2}{\alpha^2 \beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \quad (11)$$

Nota-se que B é simétrico e pode ser definido por um vetor:

$$b = \begin{bmatrix} B_{11} & B_{12} & B_{22} & B_{13} & B_{23} & B_{33} \end{bmatrix}^T \quad (12)$$

Seja o $i^{\text{ésimo}}$ vetor coluna de H . Então, temos:

$$h_i^T B h_j = v_{ij}^T b \quad (13)$$

onde

$$v_{ij} = \begin{bmatrix} h_{i1} h_{j1}, & h_{i1} h_{j2} + h_{i2} h_{j1}, & h_{i2} h_{j2}, & h_{i3} h_{j1} + h_{i1} h_{j3}, & h_{i3} h_{j2} + h_{i2} h_{j3}, & h_{i3} h_{j3} \end{bmatrix}^T \quad (14)$$

Portanto, as duas restrições 9 e 10 podem ser reescritas como 2 equações homogêneas em b , dada uma homografia:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (15)$$

Se n imagens de um plano com padrões são observados, ao empilhar n equações como 15, temos:

$$Vb = 0 \quad (16)$$

onde V é uma matriz $2n \times 6$. Se $n \geq 3$, teremos uma única solução b definida para um fator de escala. Se $n = 2$, podemos impor uma restrição de não obliquidade $\gamma = 0$.

Uma vez que b é estimado, podemos calcular todos os parâmetros da matriz intrínseca A .

4.3.3.3 Estimação de verossimilhança máxima

Em casos reais, devido a vários fatores não controláveis, as imagens capturadas de planos modelos possuem ruídos, não sendo possível solucionar a calibração da câmera pela solução fechada. Desta forma, podemos assumir que os pontos das imagens podem estar corrompidos e que os pontos de ruído são independentemente e regularmente distribuídos. A estimação máxima da verossimilhança pode ser obtida minimizando a equação:

$$\sum_{i=1}^n \sum_{j=1}^m \|m_{ij} - \hat{m}(A, R_i, t_i, M_j)\|^2 \quad (17)$$

onde $\hat{m}(A, R_i, t_i, M_j)$ é a projeção do ponto M_j na imagem i , de acordo com a equação 6. A rotação R é parametrizada por um vetor de 3 parâmetros, denotado como r , o qual é paralelo ao eixo de rotação e com magnitude igual ao ângulo de rotação. Tanto R como r são relacionados pela fórmula *Rodrigues* (MEBIUS, 2007). A minimização da equação 17 é um problema de minimização não linear, que pode ser resolvida com o algoritmo *Levenberg-Marquardt* (MARQUARDT, 1963).

4.3.3.4 Distorção provocada pelas lentes

As distorções introduzidas pelas lentes seriam possíveis de serem eliminadas se fossem utilizadas lentes parabólicas alinhadas mecanicamente com o sensor da câmera. Contudo, na prática, é muito mais fácil a fabricação de lentes esféricas, muito utilizadas nas câmeras atuais. Devido a estas limitações, há duas principais distorções: radial, provocada pela forma da lente; e a tangencial, distorção introduzida como resultado do processo de montagem.

As lentes das câmeras geralmente possuem notáveis distorções nos *pixels* localizados próximos às bordas da imagem, esta distorção é conhecida como a distorção radial da lente e dá origem ao fenômeno de efeito barril ou Olho-de-peixe. Com o uso de lentes os raios longe do centro das lentes são mais dobrados que os próximos do centro. Uma lente econômica típica, em geral, dobra os raios longe do centro de tal forma que torna perceptível o efeito barril.

No centro óptico da lente, onde está localizado o centro do sensor da câmera, a distorção é 0 e aumenta de acordo com a distância do centro da imagem. Podemos caracterizar esta distorção pelos três primeiros termos de uma expansão da série de Taylor em volta do raio

igual 0 ($r = 0$). Desta forma, podemos modelar a distorção radial de acordo com:

$$x_{r_{\text{corrigido}}} = x(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (18)$$

$$y_{r_{\text{corrigido}}} = y(1 + k_1r^2 + k_2r^4 + k_3r^6) \quad (19)$$

onde k_1 , k_2 e k_3 são os coeficientes de distorção radial da lente; (x, y) é o local original do pixel (obtidos pelo sensor da câmera) e $(x_{r_{\text{corrigido}}}, y_{r_{\text{corrigido}}})$ é a localização do pixel corrigido para a distorção radial.

A segunda distorção mais comum é a distorção tangencial que surge devido ao não paralelismo da lente com o plano de imagem durante o processo de montagem. A distorção tangencial pode ser caracterizada por dois parâmetros p_1 e p_2 , de tal forma que:

$$x_{t_{\text{corrigido}}} = x + [2p_1y + p_2(r^2 + 2x^2)] \quad (20)$$

$$y_{t_{\text{corrigido}}} = y + [p_1(r^2 + 2y^2) + 2p_2x] \quad (21)$$

onde (x, y) é a localização original do pixel; $(x_{t_{\text{corrigido}}}, y_{t_{\text{corrigido}}})$ é a localização do pixel corrigido para distorção tangencial, r é a distância do pixel do centro da imagem; p_1 e p_2 são os coeficientes de distorção tangencial.

Reorganizando as equações 18,19, 20 e 21 na forma matricial:

$$\begin{bmatrix} x_{\text{corrigido}} \\ y_{\text{corrigido}} \end{bmatrix} = (1 + k_1r^2 + k_2r^4 + k_3r^6) \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2p_1y + p_2(r^2 + 2x^2) \\ p_1(r^2 + 2y^2) + 2p_2x \end{bmatrix} \quad (22)$$

onde $(x_{\text{corrigido}}, y_{\text{corrigido}})$ é a localização do pixel corrigido para distorção radial e tangencial.

Podemos inicialmente considerar as distorções como nulas e estimar os parâmetros intrínsecos e extrínsecos através da minimização da equação 17. Então, podemos acumular os pontos e calcular os coeficientes de distorção. Uma vez estimado os coeficientes de distorção, podemos refinar a estimação dos parâmetros da câmera e das lentes ao recalculando os pontos com a equação 22 e utilizando os pontos corrigidos na minimização da equação 17, repetindo o processo até a convergência.

4.4 SEGMENTAÇÃO DAS PEÇAS

A etapa de segmentação é um importante passo para consiste em subdividir uma imagem e extrair as partes constituintes de interesse (no caso, objetos). No sistema de montagem

assistida por realidade aumentada proposto, a segmentação é necessária para o cálculo do posicionamento do objeto no plano da imagem, para o reconhecimento de objetos e na estimação da geometria do objeto.

Existem vários algoritmos de segmentação, cada um adequado a um contexto diferente. No sistema proposto, as bordas das peças são utilizadas para a segmentação. A extração de bordas é realizada por uma série de algoritmos em sequência:

1. limiar de cor;
2. detecção de bordas através *Canny Edge* (CANNY, 1986);
3. fechamento morfológico;
4. agregação das bordas em listas de pontos;
5. compactação da lista de pontos para representação de vértices.

O algoritmo de limiar de cor tem como objetivo limitar o número de partes segmentadas em apenas peças de interesse e melhorar o desempenho do sistema. A imagem é convertida do espaço RGB (*red*, *green* e *blue*) para HSV (*hue*: matiz; *saturation*: saturação; *value*: valor) e, então, apenas os *pixels* cuja matiz, saturação e valor estejam dentro dos limiares configurados são filtrados como partes de interesse para o próximo passo da segmentação. Em seguida, o *Canny Edge* é utilizado para a detecção de bordas. O fechamento morfológico com um elemento de tamanho preestabelecido é utilizado para fechar as bordas detectadas que eventualmente não estão fechadas. Após o fechamento das bordas, é necessário agregar as bordas que fazem parte de um mesmo objeto em uma estrutura de dados. Desta forma, o algoritmo de *border following* apresentado em (SUZUKI; ABE, 1985) é utilizado para transformar as bordas em uma coleção de listas de pontos. A lista de pontos de uma peça segmentada, que representa uma borda, pode apresentar pontos que pertencem a uma mesma aresta. Assim, podemos obter uma lista de vértices, a partir da lista de bordas, utilizando o algoritmo *Douglas-Peucker* (DOUGLAS; PEUCKER, 1973).

No sistema descrito, o *Canny Edge* é o principal algoritmo que permite a segmentação das peças através da detecção das bordas na imagem. Na seção a seguir é descrito o algoritmo *Canny Edge* com mais detalhes.

4.4.1 Canny Edge

O algoritmo *Canny Edge* foi desenvolvido para atingir os seguintes critérios:

- Detecção: A probabilidade de detectar pontos de arestas reais deve ser maximizado, enquanto a probabilidade de falsos positivos deve ser minimizado. Isto corresponde a maximizar a taxa de sinal-ruído.
- Localização: as arestas detectadas devem estar o mais próximo possível das arestas reais.
- Mínima resposta: Uma aresta real não deve resultar em mais de uma aresta detectada.

Há 4 passos no processo do algoritmo Canny Edge:

1. Como o detector de arestas pode ser suscetível a ruídos presentes em uma imagem não processada, um filtro Gaussiano é aplicado para reduzir o ruído;
2. O algoritmo basicamente detecta onde intensidade de cinza de uma imagem tem maior mudança. Estas áreas são descobertas com a determinação dos gradientes de uma imagem. Os gradientes de cada pixel da imagem suavizada (no passo anterior) são determinadas aplicando o operador *Sobel*. O primeiro passo é aproximar o gradiente nas direções x e y , respectivamente, aplicando os *kernels* das equações 23 e 24. As magnitudes e direções do gradiente são determinadas, respectivamente, através de $G = \sqrt{G_x^2 + G_y^2}$ e $\varphi = \arctan\left(\frac{G_y}{G_x}\right)$.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad (23)$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \quad (24)$$

3. Supressão não máxima é aplicada. Este passo remove *pixels* que não são considerados parte de uma aresta, ou seja, este passo determina se um pixel é um melhor candidato para uma aresta que seus vizinhos. Isto é realizado ao preservar todos os máximos locais da imagem de gradiente e apagar todo o resto. Para cada pixel da imagem gradiente:
 - (a) Arredondar a direção do gradiente para 0^0 , 45^0 , 90^0 ou 135^0 ;
 - (b) Comparar a resistência da aresta do pixel corrente com a resistência da aresta do pixel na direção do gradiente positivo e negativo. Isto é, se a direção do gradiente é norte ($\varphi = 90^0$), comparar com os *pixels* do norte e sul;
 - (c) Se a resistência da aresta do pixel corrente é a maior, preservar o valor da resistência da aresta. Caso contrário, suprimir o valor.

4. Histerese limiar é utilizada para determinar onde a aresta começa e termina. É utilizado dois limiares (superior e inferior):

- (a) Se um pixel gradiente é maior que o limiar superior, o pixel é aceito como uma aresta;
- (b) Se o valor do pixel gradiente é menor que o limiar inferior, o pixel é rejeitado;
- (c) Se o pixel gradiente está entre os dois limiares, será aceito como aresta somente se está conectado a um pixel que está acima do limiar superior.

4.5 REMOÇÃO DOS EFEITOS DA PERSPECTIVA

No dispositivo de assistência a câmera é posicionada de tal forma que as peças vistas estarão em perspectiva, o que gera distorções nas medidas. Uma vez que as peças estão segmentadas, é necessária a remoção do efeito da perspectiva gerada pelo posicionamento não paralelo do plano da imagem com o plano da área de trabalho. Para isto é preciso relacionar pontos do mundo real projetados em uma imagem com pontos de outra projeção de imagem.

Figura 20 - Esquema de transformação de perspectiva de um ponto em plano e projeção do ponto em planos de imagem de câmera em duas posições

Dada uma imagem de um plano obtida por uma câmera na posição p_1 , podemos estimar os pontos em uma nova imagem do mesmo plano com a câmera na posição p_2 . Seja D um ponto no plano $Z = 0$ na coordenada global, d_1 a projeção de D na imagem obtida pela câmera na posição p_1 , d_2 a projeção de D na imagem obtida pela câmera na posição p_2 . Da equação 7,

temos:

$$s_1 d_1 = H_1 D \quad (25)$$

e

$$s_2 d_2 = H_2 D \quad (26)$$

onde s_1 e s_2 são fatores de escala arbitrários; H_1 e H_2 são, respectivamente, as matrizes de homografia que relacionam os pontos d_1 e d_2 com o ponto D . Deste modo, isolando D da equação 25:

$$H_1^{-1} s_1 d_1 = D \quad (27)$$

e substituindo a equação 27 em 26:

$$s_2 d_2 = H_2 D = H_2 H_1^{-1} s_1 d_1 \quad (28)$$

Como os pontos estão definidos em coordenadas homogêneas, a relação abaixo é válida para um w arbitrário:

$$w d_2 = H_2 H_1^{-1} d_1 \quad (29)$$

Portanto, a equação 29 relaciona o ponto d_1 da projeção do ponto D , no plano da imagem com a câmera em uma posição p_1 , com o ponto d_2 da projeção do mesmo ponto D , no plano da imagem com a câmera em uma posição p_2 . Com esta relação podemos estimar pontos do objeto segmentado em planos de imagem com a câmera em outra posição para remover o efeito de perspectiva. Ainda, após o reconhecimento de objetos, a relação da equação 29 é utilizada para sobrepor a representação gráfica do modelo sobre a peça identificada na imagem capturada pela câmera.

4.6 RECONHECIMENTO DE OBJETOS

Uma vez que o sistema possui as peças segmentadas e sem efeito de perspectiva, é necessário identificar as peças de montagem. Uma vez obtidas informações sobre os objetos, o reconhecimento de objetos consiste em encontrar quais objetos e onde estão presentes em uma cena.

Em processos de montagem é comum que os fabricantes e/ou projetistas das peças tenham modelos em 2D ou 3D das peças, ou seja, poderiam ser utilizadas as informações desses modelos

para a identificação das peças. Deste modo, métodos de reconhecimento de objetos que se baseiam no conhecimento da aparência de um objeto mostram-se adequados para o uso em um sistema de montagem assistida por realidade aumentada. Estes métodos são conhecidos como reconhecimento de objetos baseado em modelos.

O reconhecimento de um objeto é realizado por correspondência de certas características de imagens e características de um modelo. Quanto às características consideradas no problema do reconhecimento de objetos, há duas principais questões a serem resolvidas: o que compõe as características e como é feita a correspondência das características. As características podem ser divididas em duas: globais, que resumem informações de porções visíveis inteiras de um objeto, e locais, que utilizam apenas informações locais, como segmentos de retas e cantos. O sistema proposto utiliza segmentos de retas para o reconhecimento de objetos através do *turning function*.

4.6.1 *Turning Function*

Um método padrão de descrição de contorno de polígono é uma lista de vértices, onde cada vértice é representado por um par de coordenadas. Os contornos segmentados das peças de montagem são descritos como listas de vértices e mudam toda vez que a peça é rotacionada ou quando a sua distância relativa à câmera é alterada, o que dificulta a comparação com outros contornos e, conseqüentemente, no reconhecimento de objetos. Portanto, as características para correspondência precisam ser invariantes à rotação, translação e escala para que seja possível a identificação das peças a partir da comparação de contornos segmentados com modelos de um banco de dados. O *turning function* apresentado em (ARKIN et al., 1991) é uma abordagem de descrição de contorno de objeto como uma representação ao qual é invariante à translação, rotação e escala. O *turning function* $\phi(s)$ mede a mudança de ângulo da tangente como função do comprimento de arco s , medido em relação a uma reta de referência que passa pelo ponto O no contorno de um polígono. Na definição do *turning function*, o perímetro do contorno é escalado de forma a ser igual a 1, ou seja, o contorno é escalado de forma que o comprimento da curva s seja definido no espaço $[0, 1]$ em $\mathbb{R}$. Assim, o *turning function* em $s = 0$ é igual ao ângulo θ_0 da tangente da aresta do contorno com uma reta de referência que passa pelo ponto O e, para um contorno fechado, o *turning function* em $s = 1$ é definido por $\phi(1) = \phi(0) + 2\pi = \theta_0 + 2\pi$. Ao seguir o caminho da curva, o *turning function* representa o ângulo formado pelo ponto em s e o ponto de referência O com a reta de referência, o ângulo é positivo se é no sentido anti-horário e é negativo se for no sentido horário. Alternativamente, o *turning function* mede o ângulo de giro acumulado ao percorrer pelo comprimento s .

Suponha um polígono representado como na figura 21a, para estimar o *turning function* da curva é necessário escalar a curva de contorno do polígono de forma que o comprimento total dessa curva seja igual a 1. A partir da curva do polígono escalado, escolhemos um ponto de início e traçamos uma reta horizontal. O ângulo θ_0 formado pela reta horizontal e a aresta inicial corresponde a $\phi(0) = \theta_0$. O gráfico da figura 21b representa o *turning function* do polígono da figura 21a.

Figura 21 - Exemplo de polígono e seu respectivo *turning function*

Suponha que para o mesmo polígono da figura 21a, escolhemos um ponto inicial diferente de modo que o deslocamento a partir de $s = 0$ seja $u \in [0, 1]$ e onde $\theta_0 = \pi$ como demonstrado na figura 22a. Ao derivar o seu *turning function* (figura 22b), podemos observar que houve um deslocamento horizontal em relação ao *turning function* da figura 21b. A escolha de ponto inicial do *turning function* corresponde a um deslocamento $u \in [0, 1]$ no comprimento da curva s , de modo que podemos representar o *turning function* como $\phi(s + u)$. Ainda, suponha que o polígono seja rotacionado, como na figura 23a. Se é assumido o mesmo ponto inicial da figura 21a ao calcular o *turning function*, como demonstrado na figura 23b, observa-se que há deslocamento vertical de valor $\frac{\pi}{2}$ em relação ao *turning function* da figura 21b. Portanto, a rotação do polígono representa um deslocamento em ϕ .

4.6.1.1 Correspondência de curvas representadas por Turning Function

A similaridade de dois polígonos é quantificada através do cálculo de distância de *turning functions* como discutido em (CAKMAKOV; CELAKOSKA, 2004). O método utilizado é denominado método poligonal. Como a escolha do ponto inicial representa um deslocamento no comprimento s do polígono e o ângulo inicial representa um deslocamento em ϕ , a distância entre dois polígonos é minimizada de acordo com a rotação e escolha de pontos de referência.

(a) Polígono com outro ponto inicial para cálculo do *turning function* (b) Gráfico do *turning function* com outro ponto inicial

Figura 22 - Exemplo de polígono com outro ponto inicial e seu respectivo gráfico de *turning function*

(a) Polígono rotacionado (b) Gráfico do *turning function* do polígono rotacionado

Figura 23 - Polígono rotacionado e seu respectivo gráfico de *turning function*

Seja dois polígonos fechados A e B , representados pelos seus respectivos *turning functions* $\phi_1(s)$ e $\phi_2(s)$, o cálculo de distância pode ser obtido por:

$$D(A, B) = \min_{\substack{\alpha \in \mathbb{R} \\ u \in [0, 1]}} \left[\int_0^1 (\phi_1(s) - \phi_2(s + u) + \rho)^2 ds \right]^{\frac{1}{2}} \quad (30)$$

onde o ângulo ótimo ρ é dado por:

$$\rho = \int_0^1 [\phi_1(s) - \phi_2(s)] ds - 2\pi u \quad (31)$$

e u representa o deslocamento a partir de um ponto inicial escolhido no *turning function*.

Para cada objeto segmentado é calculado o seu *turning function*. A distância mínima entre o contorno do objeto segmentado e cada modelo do banco de dados é calculado através da equação 30. O mínimo é atingido quando a melhor orientação entre as curvas e o ponto inicial

são selecionados. Dado um limiar de valor máximo, para cada modelo, o objeto segmentado correspondente ao modelo é aquele que possui a menor distância calculada com o modelo.

Suponha o contorno de curva polígono da figura 21a e seu respectivo contorno rotacionado em $\frac{\pi}{2}$ na figura 23a. Se calcularmos o ângulo ótimo ρ entre seus *turning functions*, representados pelas figuras 21b e 23b, concluímos que:

$$\rho = \int_0^1 [\phi_1(s) - \phi_2(s)] ds - 2\pi u = \frac{\pi}{2} - 2\pi 0 = \frac{\pi}{2} \quad (32)$$

A figura 24 representa graficamente o cálculo da equação 32. O resultado pode ser verificado ao subtrair a área do *turning function* da figura 21b pela área do *turning function* da figura 23b. A área resultante da subtração é equivalente ao ângulo ótimo ρ .

Figura 24 - Diferença da área de dois *turning functions*

No mesmo exemplo, se calcularmos a distância entre os dois polígonos através da equação 30, temos que a distância entre os dois polígonos é igual a 0, pois ambos representam o mesmo polígono rotacionado em $\frac{\pi}{2}$. Graficamente, podemos representar o cálculo da distância como sendo a área de diferença entre os *turning functions* $\phi_A(s)$ e $\phi_B(s)$ para cada ds , representado pela figura 25, menos cada área do ângulo ótimo ρ para cada ds . No exemplo, como os polígonos são exatamente iguais, para cada partição ds , temos que a subtração das áreas é igual a 0.

4.7 ASSISTÊNCIA NA MONTAGEM

Nas seções anteriores foram descritos a estrutura de dados de modelos que são utilizados para a identificação das peças, a estrutura da sequência de montagem para guiar o usuário na conclusão das tarefas de montagem, algoritmos de segmentação das peças na imagem e identificação das peças. Conforme a figura 15a, após obter cada tarefa partir da sequência de

Figura 25 - Área da diferença entre os *turning functions*

montagem, o sistema destaca a peça identificada e associada a tarefa através da sobreposição de uma representação gráfica do modelo sobre a peça. Do mesmo modo, o sistema indica ao usuário o local de encaixe da peça através de submontagens previamente estimadas durante a inicialização do *software* e valida a conexão na submontagem. Este processo é repetido para cada quadro capturado pela câmera até que é detectada a conclusão da montagem.

A seguir serão detalhados o processo de sobreposição da representação gráfica do modelo sobre a peça, a estimação das submontagens para cada tarefa e a validação das conexões na submontagem.

4.7.1 Sobreposição da representação gráfica do modelo sobre a peça

Após a segmentação e identificação das peças, o dispositivo de assistência precisa indicar ao usuário a peça relacionada a tarefa de montagem. Esta indicação é realizada através do realce da peça por sobreposição de uma representação gráfica do modelo sobre a peça identificada. A sobreposição da representação gráfica é possível por intermédio de uma transformação dos pontos do modelo na base definida na descrição para a base da imagem capturada pela câmera. Calculando-se a distância entre o centro geométrico do modelo e o centro geométrico da peça segmentada na imagem sem perspectiva podemos obter um vetor de translação $\vec{v}_d = (t_x, t_y)$. O ângulo de rotação entre a representação do modelo e a peça é o ângulo ótimo ρ obtido na identificação de objetos na equação 31 para a melhor escolha de ponto inicial. A escala η_{msp} é obtido através da relação da área do objeto segmentado e do modelo.

Assim, podemos transladar, rotacionar e escalar a representação gráfica do modelo na base m para a imagem sem efeito de perspectiva na base sp através da transformação em coorde-

nadas homogêneas:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{sp} = \begin{bmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 1 & 1 & \eta_{msp} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_m \quad (33)$$

Deste modo, ainda é necessário transformar os pontos do modelo da imagem sem perspectiva para a imagem com perspectiva para a sobreposição da representação gráfica do modelo sobre a peça. Esta transformação é realizada através da multiplicação pela inversa da matriz que transforma os pontos da imagem com perspectiva para a imagem sem perspectiva. Assim, da equação 29, para um w arbitrário da coordenada homogênea, temos que:

$$w \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_c = H_c H_{sp}^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_{sp} \quad (34)$$

onde H_c é a matriz de homografia de transformação de pontos no plano $Z = 0$ da base global para a base da imagem capturada pela câmera; H_{sp}^{-1} é a inversa da matriz de homografia de transformação de pontos do plano $Z = 0$ da base global para a base da imagem sem efeito de perspectiva.

Portanto, utilizando as equações 33 e 34 podemos transformar os pontos da representação do modelo na base m e sobrepor sobre a imagem capturada pela câmera na base c através de:

$$w \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_c = H_c H_{sp}^{-1} \begin{bmatrix} \cos\alpha & -\sin\alpha & t_x \\ \sin\alpha & \cos\alpha & t_y \\ 1 & 1 & \eta_{msp} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}_m \quad (35)$$

4.7.2 Estimação de submontagens

A indicação da posição e orientação das peças de montagem necessita do conhecimento prévio da sequência de montagem. Durante a fase de inicialização, após o carregamento dos modelos das peças e da sequência de montagem, o sistema estima a posição e orientação das peças na montagem para que seja possível guiar o usuário até que o estado final de montagem seja atingido.

Na descrição da sequência de montagem, cada tarefa é associada a uma peça de montagem e, para a peça associada são descritas as arestas que devem ser conectadas para formar uma conexão. No sistema é adotada uma base de coordenadas de referência onde todas as posições e orientação das peças são descritas na montagem. O diagrama da figura 26 apresenta como as posições e orientações das peças para cada tarefa são estimadas para a base de coordenadas de

referência da montagem. Primeiramente, a tarefa de montagem é obtida através da descrição da sequência de montagem. Em seguida, a descrição do modelo da peça associada à tarefa é obtida. Então, para pequenos incrementos de ângulo, o modelo é rotacionado até que os pares de arestas concatenadas fiquem paralelos. Após a rotação, o modelo é transladado de forma a ter o maior número de pares de arestas da conexão em contato. Caso a soma das distâncias entre as arestas da conexão seja igual a 0, a conexão é considerada como válida, os dados de orientação e posição do modelo são armazenados como uma lista de submontagens e uma nova tarefa é obtida caso não seja a última tarefa da descrição da sequência de montagem. Caso a conexão não seja considerada como válida, o ângulo de rotação é incrementado e o processo repete-se para o novo ângulo de rotação do modelo.

Figura 26 - Diagrama de blocos da estimação de submontagens

Para a primeira tarefa assume-se que a peça associada possui arestas que estão localizadas sobre o eixo x e y . Consequentemente, ao estimar cada posição e orientação da peça na base de

referência, também é estimado o estado de montagem após cada tarefa de montagem. Ou seja, para cada tarefa da sequência montagem é conhecida a submontagem resultante. No sistema, as submontagens resultantes das tarefas são acessadas posteriormente para a verificação da posição e orientação da peça para cada tarefa durante o processo de montagem.

4.7.3 Validação de conexão das peças

Uma conexão é realizada sempre que todas as arestas da tarefa de montagem estão em contato. Ao assumir que as peças de montagem são sólidas e rígidas, podemos validar a conexão da peça ao estimar sua posição e orientação relativa à submontagem. A posição e orientação da peça no estado final de montagem são estimados durante a fase de inicialização na estimação de submontagens e, durante o processo de montagem, a posição e orientação das peças são estimadas através da segmentação e reconhecimento de objetos, respectivamente. Deste modo, para validar a conexão da tarefa durante o processo de montagem é necessário verificar a posição e orientação das peças identificadas. A verificação da posição pode ser realizada calculando-se a distância entre a posição da peça na submontagem e na posição atual, enquanto a orientação pode ser verificada ao calcular a orientação relativa entre modelo e peça identificada. A conexão é considerada válida se a distância das posições e diferença das orientações forem menores que um limiar adotado.

5 IMPLEMENTAÇÃO DO SISTEMA DE MONTAGEM ASSISTIDA POR REALIDADE AUMENTADA

O protótipo do sistema foi implementado em *iPad Air 2* para auxiliar a montagem de um quebra-cabeças. Por se tratar de uma primeira iniciativa nessa área, procurou-se simplificar os algoritmos de processamento de imagens e orientação do usuário, propondo assim a resolução de um quebra-cabeças plano do tipo pentaminó (GOLOMB, 1996). O sistema implementado é denominado como MARA (Montagem Assistida por Realidade Aumentada) ou A³R (*Assembly Assisted by Augmented Reality*).

O *iPad Air 2* possui capacidade de processamento adequado para a aplicação, câmera, tela de alta resolução e sensores do tipo giroscópio e acelerômetro, além de estar amplamente disponível no mercado. Com isso os próprios sensores internos do dispositivo podem ser utilizados para obter parâmetros de orientação para a preparação das transformações de perspectiva das imagens. O dispositivo possui duas câmeras: frontal e traseira. Apenas a câmera traseira é utilizada pelo sistema, entretanto, a lente padrão do dispositivo não consegue capturar uma grande área de trabalho, por isso, uma lente grande angular (figura 27) é acoplada no dispositivo durante todo o processo de montagem.

O *software* a ser embarcado no dispositivo utiliza a biblioteca de visão computacional *OpenCV* (BRADSKI; KAEHLER, 2008) para o processamento de imagens. A biblioteca possui implementação para os seguintes algoritmos do sistema: calibração da câmera, remoção do efeito da distorção das lentes, conversão de imagens para espaço HSV, detecção de bordas através de *Canny Edge*, fechamento morfológico, vetorização de contornos por *border following* e aproximação de contorno através do algoritmo *Douglas-Peucker* (DOUGLAS; PEUCKER, 1973).

Figura 27 - Lente grande angular com distância focal de 18 mm utilizada junto com a câmera do iPad.

Os quebra-cabeças do tipo pentaminó são compostos por peças formadas por figuras geo-

métricas a partir da concatenação de 5 quadrados pelas arestas. Existem 12 tipos de peças sem contar com rotações e reflexões. As peças são identificadas de acordo com letras, conforme a figura 28a. Se espelhamos as peças, o total de peças se eleva para 18, já que as letras “T”, “V”, “I”, “X”, “U” e “W” possuem imagens espelhadas equivalentes após rotações. Um dos formatos de quebra-cabeça de pentaminó é preencher uma caixa retangular sem sobreposições e lacunas. Cada peça possui 5 unidades de quadrados, portanto, a caixa do quebra-cabeça deve possuir uma área de 60 unidades de quadrados. Os possíveis tamanhos são 6x10, 5x12, 4x15 e 3x20, sendo que cada um dos tamanhos possuem diversas soluções que já são conhecidas. A figura 28b ilustra uma das possíveis soluções para o quebra-cabeças de tamanho 6x10. Essa aplicação foi escolhida por representar um processo de montagem genérico com solução conhecida onde o usuário precisa montar as peças numa posição bem definida, a partir de uma distribuição aleatória no espaço. Além disso, o quebra-cabeças pode ser descrito através de modelos simples, o que simplifica a estrutura da solução.

(a) Denominação de cada uma das 12 peças do pentaminó.

(b) Uma das possíveis soluções do quebra-cabeças de tamanho 6x10.

Figura 28 - Pentaminós

Ao utilizar o sistema carregado no dispositivo, o *iPad* é posicionado entre o usuário e a área de trabalho, de forma que, o usuário veja as partes a serem montadas no *display* do *tablet*, conforme as figuras 29 e 30. Assim, o usuário pode ver as peças e suas mãos através do *display* juntamente com as sobreposições de informações posicionadas pelo sistema. A câmera do dispositivo captura as imagens da área de trabalho e conforme cada tarefa da montagem, a peça associada a tarefa de montagem é identificada e sua localização é indicada através de uma representação gráfica do modelo sobre a peça. Ao ser identificado a necessidade de inverter a face da peça, o sistema indica a ação através da alternância entre a representação gráfica do modelo e sua respectiva representação virada e, ao confirmar a ação um sinal sonoro é emitido. O local de conexão com as outras peças é indicada através de uma seta, cuja origem está na peça e destino no local de encaixe, e por meio de realce da posição destino.

Ainda durante o processo de montagem, uma grade de montagem de 60 quadrados, cujo

Figura 29 - Esquema da configuração da área de trabalho

Figura 30 - Foto da configuração da área de trabalho

formato representa a solução final do quebra-cabeças, é projetado sobre a imagem da tela. Assim, para guiar o usuário no posicionamento da peça na submontagem, um agrupamento de 5 quadrados que correspondem ao local de destino da peça é realçado na grade. Como parte de otimização para atingir um menor tempo de execução, a conexão é considerada válida se um número de *pixels* de cor específica localizados dentro do contorno da peça na submontagem for maior que um limiar. Deste modo, para a verificação da conexão é necessário apenas verificar a ocupação dos 5 quadrados da grade que representam o local da peça na submontagem. Como o agrupamento de 5 quadrados da grade realçados representam o contorno aproximado da peça na submontagem, a ocupação é detectada contando-se o número de *pixels* de cor específica dentro dos quadrados e comparando a totalização com um limiar. As cores dos *pixels* na verificação da conexão de montagem são previamente calibradas para as cores das

peças antes do início do processo de montagem. Para sinalizar a confirmação da conexão da peça à submontagem, a cor de realce da grade de montagem é transformada e, do mesmo modo que a confirmação da ação de inversão da face da peça, um som é emitido ao ser confirmado a conexão. O sistema detecta que a montagem foi concluída ao finalizar a última tarefa de montagem e exibe a mensagem de conclusão para o usuário.

Nas seções a seguir serão descritos a configuração necessária para o uso do dispositivo de assistência, os arquivos de descrição de modelos e sequência de montagem e como o efeito da perspectiva é removido utilizando os sensores do dispositivo.

5.1 CONFIGURAÇÃO DO SISTEMA

Devido a variações de ambiente e possível uso de peças de pentaminó de diferentes cores, o sistema foi implementado de forma a ser possível o ajuste de parâmetros de processamento de imagens.

Os parâmetros dos algoritmos de detecção de bordas *Canny Edge*, fechamento morfológico, aproximação de contorno e cálculo de distância entre *turning functions* podem ser configurados através de uma tela de configuração conforme a figura 31. Na tela, o campo *Canny edge lower threshold* é utilizado para configurar o limiar inferior da histerese. O limiar superior da histerese é estimado como sendo 3 vezes o limiar inferior definido. O tamanho do elemento quadrado a ser utilizado no fechamento morfológico é definido no campo *morphology element size*, sendo que o valor define o número de linhas e colunas do elemento. Para a aproximação de contorno é definido um valor precisão de aproximação em frações de 0 a 1 no campo *contour approximation precision*, cujo valor 1 significa que o contorno aproximado deve ser exatamente igual ao contorno original. O campo *turning function minimum distance* é utilizado para definir o valor mínimo de distância na estimação da similaridade. *Turning functions* de modelos de cuja distância com os *turning functions* das peças segmentadas são menores que o valor configurado não são considerados como candidatos na identificação da peça. Nesta tela ainda é possível iniciar o processo de calibração da câmera e configuração dos limiares de cor para segmentação e detecção de posicionamento das peças na grade de montagem.

Ao iniciar a configuração dos limiares de cor a tela da figura 32 é exibida. A configuração é realizada deslizando os *slides* de limiares inferior e superior para a matiz (*hue*), saturação (*saturation*) e valor (*value*). Nesta tela de configuração é mostrado ao usuário a imagem após aplicar o limiar de cor para facilitar a configuração dos limiares.

Antes de iniciar o processo de montagem é necessário calibrar a câmera para que sejam removidos os efeitos da distorção das lentes e para calcular os parâmetros intrínsecos da câmera. A calibração é necessária na primeira execução do aplicativo no dispositivo ou se

Figura 31 - Tela de configuração de parâmetros de algoritmos de segmentação e reconhecimento de objetos

houver a troca da lente acoplada na câmera. Na seção a seguir é descrito o método de calibração da câmera.

5.1.1 Calibração da câmera

A calibração da câmera é realizada utilizando o método de Zhang (1999) descrito na seção 4.3.3. Para a calibração, deve-se relacionar pontos de um padrão plano com as imagens. O *OpenCV* já possui funções para estimar os parâmetros intrínsecos e coeficientes de distorção. Para sua utilização é necessário utilizar como entrada imagens de padrão de geometria conhecida. O padrão de geometria conhecida utilizado é um xadrez de tamanho 9x6, conforme a figura 33. Ao iniciar o processo de calibração, as imagens são obtidas automaticamente em intervalos de 1s, pelo menos 12 imagens do padrão xadrez são obtidas com a câmera do dispositivo em diversas posições e orientações diferentes. Algumas imagens obtidas durante este processo não são adequadas para a detecção do padrão de xadrez, portanto, o sistema não os utilizam na contagem de imagens obtidas para serem utilizados na calibração. Como resultado é obtido uma matriz A de tamanho 3x3 que descreve os parâmetros intrínsecos da câmera, um vetor de coeficientes de distorção das lentes $[k_1 \ k_2 \ p_1 \ p_2 \ k_3]$ e o erro de reprojeção. O software apenas considera os parâmetros como corretos se o erro de reprojeção

Figura 32 - Tela de configuração de limiares de cor

é menor que 0,5. Caso o erro de reprojeção seja menor que o definido, é necessário repetir o processo de calibração até que se obtenha um erro de reprojeção adequado. Após a calibração da câmera, o efeito de distorção das lentes é automaticamente corrigido em todas as imagens capturadas.

5.2 DESCRIÇÃO DE MODELOS E SEQUÊNCIA DE MONTAGEM ATRAVÉS DE ARQUIVOS

Os modelos das peças de pentaminó e sequência de montagem são carregados a partir de arquivo em formato JSON seguindo a estrutura descrita nas seções 4.1 e 4.2.

Para a descrição dos modelos, os rótulos adotados para as peças adotados são os mesmos da denominação das peças de pentaminó da figura 28a. Para cada modelo de peça adotou-se uma base para descrição dos vértices onde pelo menos uma aresta está localizada em cada eixo. A listagem 1 demonstra uma parte da descrição dos modelos das peças onde é descrito o modelo da peça "V" do pentaminó. No exemplo, o rótulo das arestas é representado pela concatenação do rótulo da peça com um sequencial. A listagem 2 apresenta parte do mesmo arquivo em JSON da listagem 1 para a descrição do modelo da peça "U".

Na descrição da sequência de montagem, cada elemento da lista representa uma tarefa de


```

{
  "Parts":
  [
    ...
    {
      "Label": "V",
      "Edges":
      [
        {
          "Label": "V0",
          "Vertices": [{"X": 0,"Y": 0},{X": 0,"Y": 300 }]
        },
        {
          "Label": "V1",
          "Vertices": [{"X": 0,"Y": 300},{X": 100,"Y": 300 }]
        },
        {
          "Label": "V2",
          "Vertices": [{"X": 100,"Y": 300},{X": 100,"Y": 100 }]
        },
        {
          "Label": "V3",
          "Vertices": [{"X": 100,"Y": 100},{X": 300,"Y": 100 }]
        },
        {
          "Label": "V4",
          "Vertices": [{"X": 300,"Y": 100},{X": 300,"Y": 0 }]
        },
        {
          "Label": "V5",
          "Vertices": [{"X": 300,"Y": 0},{X": 0,"Y": 0 }]
        }
      ]
    },
    ...
  ]
}

```

Listagem 1 - Exemplo de parte do arquivo JSON com descrição do modelo da peça "V"

```

{
  "Parts":
  [
    ...
    {
      "Label": "U",
      "Edges":
      [
        {
          "Label": "U0",
          "Vertices": [{"X": 0,"Y": 0}, {"X": 0,"Y": 100}]
        },
        {
          "Label": "U1",
          "Vertices": [{"X": 0,"Y": 100}, {"X": 100,"Y": 100}]
        },
        {
          "Label": "U2",
          "Vertices": [{"X": 100,"Y": 100}, {"X": 100,"Y": 200}]
        },
        {
          "Label": "U3",
          "Vertices": [{"X": 100,"Y": 200}, {"X": 0,"Y": 200}]
        },
        {
          "Label": "U4",
          "Vertices": [{"X": 0,"Y": 200}, {"X": 0,"Y": 300}]
        },
        {
          "Label": "U5",
          "Vertices": [{"X": 0,"Y": 300}, {"X": 200,"Y": 300}]
        },
        {
          "Label": "U6",
          "Vertices": [{"X": 200,"Y": 300}, {"X": 200,"Y": 0}]
        },
        {
          "Label": "U7",
          "Vertices": [{"X": 200,"Y": 0}, {"X": 0,"Y": 0}]
        }
      ]
    },
    ...
  ]
}

```

Listagem 2 - Exemplo de parte do arquivo JSON com descrição do modelo da peça "U"

Figura 33 - Padrão tamanho 9 x 6 utilizado na calibração

montagem. Para cada tarefa de montagem, a peça associada (*part*) é identificada pelo seu rótulo e as conexões (*connections*) de arestas da peça com a submontagem são descritas como uma lista de pares de rótulos que identificam as arestas (*edges*). A listagem 3 descreve uma parte da descrição da sequência de montagem. No exemplo são descritas a primeira, segunda, terceira, quarta, quinta e sexta tarefa de montagem da sequência de montagem. Nas tarefas são movimentadas e encaixadas as peças "W", "T", "L", "I", "Z" e "Y" sequencialmente. Para a primeira tarefa, assume-se que duas arestas da peça associada a tarefa devem estar no eixo x e y da base de coordenadas da montagem. No exemplo, as arestas "W2" e "W3" estão no eixo x e y , respectivamente. Na segunda tarefa, representado pela sequência (*sequence*) "1", a peça "T" deve ser movida de modo que as arestas "T0", "T1" e "T2" sejam concatenadas às arestas "W1", "W0" e "W9", nesta ordem. Do mesmo modo, na terceira tarefa, a peça "L" deve ser conectada de modo que as arestas "L4" e "L5" estejam em contato com as respectivas arestas "W5" e "W4". Sucessivamente as tarefas são descritas deste modo até a última tarefa de montagem. Observa-se que a cada tarefa da sequência a peça associada é conectada a uma peça já pertencente à submontagem. Assim, o sistema pode estimar as submontagens sequencialmente após carregar o arquivo JSON.

No sistema implementado alguns arquivos com sequências de montagem de soluções diferentes do quebra-cabeça já são pré-carregados e são selecionados a partir de um menu, conforme a figura 34.

5.3 REMOÇÃO DOS EFEITOS DA PERSPECTIVA

Antes do reconhecimento das peças planas, um passo anterior é transformar a perspectiva para a visão de uma câmera com plano de imagem paralelo ao plano das peças. Esta é uma

```

{
  "AssemblySequence": [
    {
      "Sequence": "0",
      "Part": "W",
      "Connections": [{ "Edges": [ "W2" ] }, { "Edges": [ "W3" ] } ]
    },
    {
      "Sequence": "1",
      "Part": "T",
      "Connections": [{ "Edges": [ "T0", "W1" ] }, { "Edges": [ "T1", "W0" ] }, { "Edges": [ "T2", "W9" ] } ]
    },
    {
      "Sequence": "2",
      "Part": "L",
      "Connections": [{ "Edges": [ "L4", "W5" ] }, { "Edges": [ "L5", "W4" ] } ]
    },
    {
      "Sequence": "3",
      "Part": "I",
      "Connections": [{ "Edges": [ "I1", "T4" ] }, { "Edges": [ "I0", "T5" ] } ]
    },
    {
      "Sequence": "4",
      "Part": "Z",
      "Connections": [{ "Edges": [ "Z4", "W6" ] }, { "Edges": [ "Z1", "I1" ] } ]
    },
    {
      "Sequence": "5",
      "Part": "Y",
      "Connections": [{ "Edges": [ "Y1", "I1" ] }, { "Edges": [ "Y2", "Z0" ] } ]
    },
    ...
  ]
}

```

Listagem 3 - Exemplo de descrição de sequência de montagem formatada em JSON

Figura 34 - Tela de seleção de sequência de montagem carregada no sistema

importante fase para simplificar a quantidade de informação necessária para a identificação da peça, uma vez que distorções de perspectiva torna a identificação de peças em uma tarefa mais complexa.

Seja a figura 35, através da relação do triângulo formado pela origem do sistema de coordenada da câmera C , origem do sistema de coordenada global O e com o ponto de contato do *tablet* com a superfície T , temos que:

$$a = h \cdot \sin(\theta) \quad (36)$$

e

$$b = a \cdot \tan \theta \quad (37)$$

Assim, substituindo a equação 36 em 37, temos que:

$$b = h \cdot \frac{\sin^2(\theta)}{\cos(\theta)} \quad (38)$$

Da definição da homografia da equação 6, podemos definir a matriz de homografia H_c que relaciona os pontos do plano $X_g Y_g$ da coordenada global com o plano da imagem capturada

Figura 35 - Esquema de arranjo do *tablet* e sistema de coordenadas da câmera, global e da câmera virtual

através de:

$$A \begin{bmatrix} -\cos(\theta) & 0 & b \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & a \end{bmatrix} = A \begin{bmatrix} -\cos(\theta) & 0 & h \cdot \frac{\sin^2(\theta)}{\cos(\theta)} \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & h \cdot \sin(\theta) \end{bmatrix} = H_c \quad (39)$$

onde A é a matriz de parâmetros intrínsecos da câmera.

Para obter os pontos do plano X_gY_g sem efeitos da perspectiva devemos simular uma câmera virtual cuja origem dos eixos esteja em $C_v = (0, 0, h_v)_G$. Desta maneira, a matriz de homografia H_v que relaciona os pontos do plano X_gY_g com o plano da imagem da câmera em C_v é:

$$A_v \begin{bmatrix} \cos(\pi) & 0 & 0 \\ 0 & 1 & 0 \\ -\sin(\pi) & 0 & h_v \end{bmatrix} = A_v \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & h_v \end{bmatrix} = H_v \quad (40)$$

onde A_v é a matriz de parâmetros intrínsecos da câmera virtual.

Na implementação do sistema, a matriz de parâmetros intrínsecos da câmera virtual é considerada como igual a matriz de parâmetros intrínsecos estimado na calibração da câmera.

Assim, utilizando a relação da equação 29, temos que:

$$w d_v = H_v H_c^{-1} d_c = A \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & h_v \end{bmatrix} A^{-1} \begin{bmatrix} -\cos(\theta) & 0 & h \cdot \frac{\sin^2(\theta)}{\cos(\theta)} \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & h \cdot \sin(\theta) \end{bmatrix}^{-1} d_c \quad (41)$$

onde w é o fator de escala arbitrário da transformação homogênea; d_v é o ponto no plano

de imagem da câmera virtual; h_v é distância do plano da imagem da câmera virtual e o plano $X_g Y_g$ da coordenada global; A é a matriz de parâmetros intrínsecos da câmera; θ é o ângulo de inclinação do *tablet*; d_c é o ponto projetado no plano da imagem da câmera.

O *tablet iPad Air 2* possui sensores do tipo giroscópio e acelerômetro, ambos os sensores foram utilizados para a obtenção do ângulo θ de inclinação do *tablet* através de *API* oferecida pelo sistema operacional do *tablet*. Além disso, a medida h foi obtida através de medição direta. Deste modo, podemos utilizar a equação 41 para remover o efeito da perspectiva da imagem das peças capturada pela câmera do dispositivo.

6 RESULTADOS E ANÁLISES DO SISTEMA DE MONTAGEM ASSISTIDA POR REALIDADE AUMENTADA

6.1 CALIBRAÇÃO DA CÂMERA

Para obter os resultados da calibração da câmera, o processo foi repetido 12 vezes com a lente grande angular da figura 27 acoplada. No processo com o menor erro de reprojeção (erro de reprojeção igual a 0,266) foram estimados os seguintes parâmetros:

$$A = \begin{bmatrix} 576 & 0 & 342 \\ 0 & 581 & 214 \\ 0 & 0 & 1 \end{bmatrix} \quad (42)$$

$$d = \begin{bmatrix} -0,232 & 0,130 & 0,00176 & -0,0115 & -0,155 \end{bmatrix} \quad (43)$$

As figuras 36, 37, 38 e 39 apresentam imagens exemplos obtidas com os parâmetros de calibração acima. Observa-se o efeito das distorções radiais provocadas pela lente grande angular nas figuras 36a, 37a, 38a e 39a com imagens obtidas pela câmera sem processamento. Nas figuras 36b, 37b, 38b e 39b são apresentadas imagens com correção de distorções das lentes para o caso 1, 2, 3 e 4, respectivamente. Observa-se que na figura 40, onde é apresentado um exemplo com mais detalhes, o contraste diminui a medida que os pontos são distanciados do centro. Este efeito influencia na segmentação, uma vez que a segmentação depende do contraste da imagem.

(a) Imagem obtida pela câmera sem processamento

(b) Imagem corrigida para distorção das lentes

Figura 36 - Caso 1 - Resultado da remoção de distorção das lentes

(a) Imagem obtida pela câmera sem processamento

(b) Imagem corrigida para distorção das lentes

Figura 37 - Caso 2 - Resultado da remoção de distorção das lentes

(a) Imagem obtida pela câmera sem processamento

(b) Imagem corrigida para distorção das lentes

Figura 38 - Caso 3 - Resultado da remoção de distorção das lentes

(a) Imagem obtida pela câmera sem processamento

(b) Imagem corrigida para distorção das lentes

Figura 39 - Caso 4 - Resultado da remoção de distorção das lentes

Figura 40 - Imagem sem efeito de distorção das lentes. Observa-se diminuição do contraste próximo das bordas

6.2 SEGMENTAÇÃO DAS PEÇAS

Para avaliar a segmentação das peças, 35 imagens com as 12 peças do pentaminó em orientações e posições aleatórias foram processadas, ou seja, foram obtidas 35 imagens de cada peça, totalizando 420 peças. As imagens processadas foram obtidas em um mesmo ambiente, sem variações de luz com o tablet fixo em uma única posição. Os parâmetros dos algoritmos da segmentação foram selecionados de modo que o maior número de peças fossem segmentadas. Os seguintes parâmetros foram utilizados no teste de segmentação:

- Fechamento morfológico com elemento de tamanho 2x2;
- Limiar inferior de histerese do *Canny edge*: 50;
- Limiar superior de histerese do *Canny edge*: 150;
- Precisão da aproximação de contorno: 1,5%.

Analisando as peças que não puderam ser segmentadas, notou-se que a segmentação não foi possível devido a bordas não fechadas e devido ao efeito mais agressivo da correção da distorção radial nas peças localizadas próximas às bordas das imagens. Os parâmetros de fechamento morfológico utilizados não foram suficientes para fechar algumas bordas, deste modo, o algoritmo de *border following* acumulou os pontos da borda como uma borda aberta. A taxa de segmentação pode ter um melhor desempenho se um elemento de tamanho maior

Tabela 6 - Resultado da segmentação

Peça	Número de imagens da peça	Número de peças segmentadas	Taxa de segmentação
F	35	34	0,97
I	35	35	1
L	35	34	0,97
N	35	34	0,97
P	35	35	1
T	35	35	1
U	35	35	1
V	35	35	1
W	35	34	0,97
X	35	35	1
Y	35	35	1
Z	35	32	0,91
<i>Total</i>	<i>420</i>	<i>413</i>	<i>0,98</i>

for utilizado no fechamento morfológico, contudo, elementos maiores podem deformar a forma das peças, o qual dificulta a tarefa de reconhecimento de objetos.

A figura 41 apresenta os resultados de segmentação para os 4 casos apresentados na calibração.

6.3 REMOÇÃO DO EFEITO DE PERSPECTIVA

Para a remoção do efeito de perspectiva foi utilizada a equação 41. Os seguintes parâmetros foram utilizados:

- Distância do ponto de contato T do *tablet* com a base da câmera C : $h = 160mm$;
- Distância do plano de imagem da câmera virtual com o plano X_gY_g : $h_v = 700mm$.

O ângulo de inclinação θ foi obtido automaticamente pelos sensores do *tablet*. A figura 42 apresenta exemplos de imagens sem o efeito de perspectiva dos mesmos casos obtidos na calibração e segmentação.

(a) Resultado de segmentação do caso 1

(b) Resultado de segmentação do caso 2

(c) Resultado de segmentação do caso 3

(d) Resultado de segmentação do caso 4

Figura 41 - Imagens exemplo de segmentação

6.4 RECONHECIMENTO DAS PEÇAS DE MONTAGEM

Testes foram realizados para avaliar o algoritmo de reconhecimento de objetos utilizado no sistema implementado. Durante o processo de reconhecimento de objetos, assume-se que as peças de pentaminó são únicas e apenas partes segmentadas cujo *turning function* possuem distâncias menores que 0,8 com os modelos são consideradas como candidato para identificação. Como este processo depende diretamente do resultado da segmentação, utilizamos apenas as partes que foram segmentadas nos testes. Os resultados alcançados são apresentados na tabela 7. De um total de 413 partes segmentadas, 407 peças foram identificadas corretamente.

(a) Resultado de remoção do efeito de perspectiva para o caso 1

(b) Resultado de remoção do efeito de perspectiva para o caso 2

(c) Resultado de remoção do efeito de perspectiva para o caso 3

(d) Resultado de remoção do efeito de perspectiva para o caso 4

Figura 42 - Imagens exemplo da remoção do efeito de perspectiva

Tabela 7 - Resultado do reconhecimento de objetos

Peça	Número de partes segmentadas	Quantidade de peças identificadas	Taxa de identificação
F	34	33	0,97
I	35	35	1
L	34	33	0,97
N	34	34	1
P	35	34	0,97
T	35	35	1
U	35	33	0,94
V	35	35	1
W	34	34	1
X	35	35	1
Y	35	34	0,97
Z	32	32	1
<i>Total</i>	<i>413</i>	<i>407</i>	<i>0,98</i>

A figura 43 apresenta resultados da identificação para os casos 1, 2, 3 e 4, listados anteriormente nos resultados da calibração, segmentação e remoção do efeito de perspectiva. Nos resultados, as representações gráficas dos modelos estão desenhadas sobre as suas respectivas peças de correspondência.

(a) Resultado do reconhecimento de objetos do caso 1

(b) Resultado do reconhecimento de objetos do caso 2

(c) Resultado do reconhecimento de objetos do caso 3

(d) Resultado do reconhecimento de objetos do caso 4

Figura 43 - Imagens exemplo de reconhecimento das peças do pentaminó

Analisando as imagens dos resultados, notou-se que algumas peças próximas às bordas das imagens não foram identificadas. Estas peças tiveram a sua forma afetada pela distorção radial das lentes. A figura 44 apresenta um caso de falha na identificação da peça “U”. Nota-se na figura 44 que a parte côncava da peça “U” encontra-se deformada.

6.5 MONTAGEM ASSISTIDA

O sistema guia o usuário através de símbolos gráficos e sonoros. Quando é necessário inverter a face de uma peça, o sistema exibe alternadamente duas sobreposições da representação do modelo sobre a peça, uma em cor branca e outra em cor verde. A representação em cor branca representa o estado atual da peça enquanto a cor verde representa o estado da peça após ser virada. Quando é detectado que a peça foi virada, um alerta sonoro é emitido. Exemplos do símbolo são representados nas figuras 45, 46, 47, 48, 49 e 50. Em cada uma das figuras é

(a) Imagem corrigida para distorção das lentes

(b) Imagem da segmentação

(c) Imagem sem efeito de perspectiva

(d) Resultado da identificação - peça "U" não foi identificada

Figura 44 - Caso em que a peça "U" não foi identificada devido à deformação de sua forma

apresentado um exemplo do símbolo da ação de inverter a face da peça para cada peça que possui espelhamento: peças "F", "L", "N", "P", "Y" e "Z".

(a) Peça "F" detectada em branco

(b) Símbolo indicando para inverter a face da peça "F"

Figura 45 - Símbolo de ação de inversão da face da peça "F"

(a) Peça "L" detectada em branco

(b) Símbolo indicando para inverter a face da peça "L"

Figura 46 - Símbolo de ação de inversão da face da peça "L"

(a) Peça "N" detectada em branco

(b) Símbolo indicando para inverter a face da peça "N"

Figura 47 - Símbolo de ação de inversão da face da peça "N"

(a) Peça "P" detectada em branco

(b) Símbolo indicando para inverter a face da peça "P"

Figura 48 - Símbolo de ação de inversão da face da peça "P"

(a) Peça "Y" detectada em branco

(b) Símbolo indicando para inverter a face da peça "Y"

Figura 49 - Símbolo de ação de inversão da face da peça "Y"

(a) Peça "Z" detectada em branco

(b) Símbolo indicando para inverter a face da peça "Z"

Figura 50 - Símbolo de ação de inversão da face da peça "Z"

A figura 51 apresenta um estado intermediário de um processo de montagem guiado. As peças que já foram movidas e conectadas em tarefas anteriores estão dispostas sob a grade onde as células estão em cor amarela, representando a submontagem. A peça de montagem a ser movida na tarefa da sequência está realçada em cor verde e uma seta indica ao usuário onde a peça deve ser movida para ser conectada na submontagem. As células em realce de cor vermelha na grade indicam a posição e orientação de encaixe da peça de montagem. O realce em cor vermelha na grade onde a peça deve ser colocada é exibido mesmo quando a peça não está sendo detectada na imagem. Assim, mesmo que a peça não seja reconhecida no quadro de imagem corrente, o processo de montagem não é interrompido, pois o usuário pode inferir qual a peça da tarefa pelo formato apresentado na grade e movê-la para outro local da área de trabalho para que o sistema a reconheça. Quando é detectado que a peça está no local indicado da grade, um som de confirmação audível é produzido. Uma nova tarefa é obtida após a confirmação e o processo é repetido até que seja detectado a conclusão da montagem, como pode ser observado na figura 51b.

(a) Processo de montagem em estado intermediário

(b) Processo de montagem concluído

Figura 51 - Imagens do processo de montagem assistida

No sistema foram carregados 5 descrições diferentes de sequência de montagem: duas soluções para o quebra-cabeças de tamanho 6×10 e uma solução para cada quebra-cabeças de tamanho 5×12 , 4×15 e 3×20 . Deste modo, o sistema foi validado para cada tamanho possível para o quebra-cabeças do tipo pentaminó. Cada uma das figuras 52 e 53 apresenta a montagem completa de uma solução para o quebra-cabeças de tamanho 6×10 e as figuras 54, 55 e 56 apresentam a montagem das soluções de quebra-cabeças de tamanho 5×12 , 4×15 e 3×20 , respectivamente.

6.6 TEMPO DE EXECUÇÃO DOS ALGORITMOS

Execução em tempo real é desejável em sistemas de realidade aumentada, portanto, é importante avaliar o tempo de execução de cada algoritmo de processamento de imagens e reconhecimento de objetos. Para avaliação do tempo de execução dos processos, peças do pentaminó foram dispostas de modo que a câmera pudesse capturar todas as 12 peças e, então, os algoritmos de processamento de imagens foram executados para 500 quadros de imagens. A tabela 8 resume os resultados do tempo de execução de cada algoritmo. O tempo de execução total foi de cerca de 109 ms, em outras palavras, a taxa de atualização de imagens foi de aproximadamente 10 quadros por segundo (10 fps). 10 fps são suficientes para tempo real nesta aplicação, pois o processamento de imagens é feito em paralelo à exibição de imagens capturadas pela câmera. Assim, o usuário vê a imagem efetivamente a 30 fps e mesmo com a atualização da Realidade Aumentada a 10 fps o usuário não percebe nenhuma tremulação na imagem.

O processo de segmentação por ser composto por vários algoritmos, representa o maior

Figura 52 - Montagem de uma solução do quebra-cabeças de tamanho 6x10

Figura 53 - Montagem de uma segunda solução do quebra-cabeças de tamanho 6x10

Figura 54 - Montagem de um quebra-cabeças de tamanho 5x12

Figura 55 - Montagem de um quebra-cabeças de tamanho 4x15

Figura 56 - Montagem de um quebra-cabeças de tamanho 3x20

tempo de execução: cerca de 63% do tempo de execução total. O processo de reconhecimento de objetos e a remoção do efeito de distorção das lentes são o segundo processo com maior tempo de execução. Como a derivação do *turning function* das peças segmentadas e seu cálculo de distância com os modelos das peças podem ser paralelizados, o processo de reconhecimento de imagens pode ter seu desempenho otimizado em um trabalho decorrente.

Tabela 8 - Tempo de execução dos processos

Processo	Tempo (ms)	
	Média	Desvio padrão
Remoção do efeito de distorção das lentes	19,1	2,18
Segmentação	68,9	4,39
Remoção do efeito de perspectiva	0,435	0,0706
Reconhecimento de objetos	20,7	1,71
<i>Total</i>	<i>109</i>	<i>8,35</i>

7 CONCLUSÃO

Neste trabalho foi descrito um sistema para assistência em processos de montagem com o uso de realidade aumentada. O sistema foi concebido para detectar as peças e auxiliar a montagem através de processamento de imagens sem que se necessite marcadores para rastreamento e identificação das peças, dispensando uma preparação inicial do ambiente de trabalho. Além disso, o sistema foi concebido em forma de *software* de modo a ser instalado e executado embarcado em dispositivos *off-the-shelf* sem a necessidade de uso de equipamento customizado ou comunicação com outros equipamentos para o processamento de imagens. Como consequência, espera-se que o uso da tecnologia para auxiliar montagens seja popularizado.

Para a validação do sistema descrito, um protótipo foi implementado para guiar a montagem de um quebra-cabeças plano do tipo pentaminó. Foram obtidos resultados para cada etapa de processamento de imagens: remoção do efeito de distorção das lentes, segmentação, remoção do efeito de perspectiva e reconhecimento de objetos. Nos testes, o processo de segmentação conseguiu segmentar cerca de 98% dos objetos de interesse. O reconhecimento de objetos foi realizado através de um algoritmo que utiliza o conceito do *turning function* para o cálculo de similaridade entre as peças e os objetos. O algoritmo de reconhecimento de objetos identificou cerca de 98% das peças segmentadas. Apesar de algumas peças não serem segmentadas ou reconhecidas em algumas situações, principalmente quando se encontram nas bordas da imagem, o processo de guia não é interrompido, pois o sistema indica qual o formato da peça através de realce de cor na grade de montagem e, assim, o usuário pode deduzir qual é a peça da tarefa. O processo de guia de montagem foi validado para pelos menos uma solução de cada um dos 4 diferentes tamanhos possíveis de quebra-cabeças do tipo pentaminó: 6x10, 5x12, 4x15 e 3x20. Foram apresentadas imagens sequenciais da montagem completa de cada uma das soluções. Também avaliou-se o tempo de execução das etapas de processamento de imagens e, para cada quadro de imagem o tempo de execução total foi de 109 ms, em outras palavras, foi alcançado uma taxa de atualização de quadros de 10 fps. A exibição das imagens e o processamento de imagens são executados em paralelo, portanto, para o usuário do sistema não há atraso na exibição das imagens devido ao processamento de imagens. Desta modo, 10 fps são suficientes para tempo real nesta aplicação.

Como trabalhos decorrentes foram geradas duas publicações: um artigo em conferência (NISHIHARA; OKAMOTO, 2015) sobre o sistema de reconhecimento de objetos e um capítulo de livro (OKAMOTO; NISHIHARA, 2016) sobre o sistema MARA (ou A³R). Há espaço para melhorias no desempenho do reconhecimento de objetos, como a paralelização do cálculo de *turning function* das peças segmentadas e o cálculo de distância entre peças segmentada e modelos. Assim como outros trabalhos (TANG et al., 2003; WIEDENMAIER et al., 2003;

BAIRD; BARFIELD, 1999), pode-se comparar o tempo de montagem de usuários com o sistema implementado e outros meios, como o manual em papel ou digital. Ainda, é interessante comparar o desempenho do sistema para diferentes condições de uso, como a utilização em ambiente sem iluminação controlada. Por fim, poderia ser implementado um sistema de rastreamento para o sistema, o que possibilitaria o uso do sistema sem que o tablet esteja fixo e possibilitaria a implementação em uma plataforma do tipo *Hololens*.

Referências

ALVAREZ, H.; AGUINAGA, I.; BORRO, D. Providing guidance for maintenance operations using automatic markerless Augmented Reality system. **2011 10th IEEE International Symposium on Mixed and Augmented Reality**, IEEE, p. 181–190, out. 2011. DOI: 10.1109/ISMAR.2011.6092385.

ARKIN, E. M.; CHEW, L. P.; HUTTENLOCHER, D. P.; KEDEM, K.; MITCHELL, J. S. B. An efficiently computable metric for comparing polygonal shapes. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 13, n. 3, p. 209–216, mar. 1991. DOI: 10.1109/34.75509.

AZUMA, R.; BAILLOT, Y.; BEHRINGER, R.; FEINER, S.; JULIER, S.; MACINTYRE, B. Recent advances in augmented reality. **IEEE Computer Graphics and Applications**, v. 21, n. 6, p. 34–47, 2001. DOI: 10.1109/38.963459.

AZUMA, R. T. A Survey of Augmented Reality. **Teleoperators and Virtual Environments**, Citeseer, v. 6, n. 4, p. 355–385, 1997.

BAIRD, K. M.; BARFIELD, W. Evaluating the effectiveness of augmented reality displays for a manual assembly task. **Virtual Reality**, v. 4, n. 4, p. 250–259, dez. 1999. DOI: 10.1007/BF01421808.

BARANDIARAN, J.; BORRO, D. Edge-Based Markerless 3D Tracking of Rigid Objects. **17th International Conference on Artificial Reality and Telexistence (ICAT 2007)**, IEEE, p. 282–283, nov. 2007. DOI: 10.1109/ICAT.2007.62.

BOUD, A.; HANIFF, D.; BABER, C.; STEINER, S. Virtual reality and augmented reality as a training tool for assembly tasks. **1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)**. p. 32–36, 1999. DOI: 10.1109/IV.1999.781532.

BRADSKI, G.; KAEHLER, A. **Learning OpenCV: Computer Vision with the OpenCV Library**. 1. ed. [S.l.]: O'Reilly Media, 2008. 555 p.

CAKMAKOV, D.; CELAKOSKA, E. Estimation of curve similarity using turning functions. **International Journal of Applied Mathematics**, v. 15, p. 403–416, 2004.

CANNY, J. A computational approach to edge detection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 8, n. 6, p. 679–698, jun. 1986.

CARMIGNIANI, J.; FURHT, B.; ANISETTI, M.; CERAVOLO, P.; DAMIANI, E.; IVKOVIC, M. Augmented reality technologies, systems and applications. **Multimedia Tools and Applications**, Springer Netherlands, v. 51, n. 1, p. 341–377, dez. 2010. DOI: 10.1007/s11042-010-0660-6.

CAUDELL, T.; MIZELL, D. Augmented reality: an application of heads-up display technology to manual manufacturing processes. **Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences**, IEEE, p. 659–669 vol.2, 1992. DOI: 10.1109/HICSS.1992.183317.

CHEN, H.; LEE, A. S.; SWIFT, M.; TANG, J. C. 3d collaboration method over hololens™ and skype™ endpoints. **Proceedings of the 3rd International Workshop on Immersive Media Experiences**. p. 27–30, 2015. DOI: 10.1145/2814347.2814350.

CHI, H.-L.; KANG, S.-C.; WANG, X. Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. **Automation in Construction**, Elsevier B.V., jan. 2013. DOI: 10.1016/j.autcon.2012.12.017.

CHIANG, H.-K.; CHOU, Y.-Y.; CHANG, L.-C.; HUANG, C.-Y.; KUO, F.-L.; CHEN, H.-W. An augmented reality learning space for PC DIY. **Proceedings of the 2nd Augmented Human International Conference on - AH '11**, New York, New York, USA, ACM Press, n. 2, p. 1–4, 2011. DOI: 10.1145/1959826.1959838.

DE MELLO, L. S. H.; SANDERSON, A. C. Representations of mechanical assembly sequences. **IEEE Transactions on Robotics and Automation**, v. 7, n. 2, p. 211–227, 1991. DOI: 10.1109/70.75904.

DOUGLAS, D. H.; PEUCKER, T. K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. **Cartographica: The International Journal for Geographic Information and Geovisualization**, v. 10, n. 2, p. 112–122, 1973. DOI: 10.3138/FM57-6770-U75U-7727.

FEINER, S.; MACINTYRE, B.; SELIGMANN, D. Knowledge-based augmented reality. **Communications of the ACM**, v. 36, n. 7, p. 53–62, jul. 1993. DOI: 10.1145/159544.159587.

FIORENTINO, M.; UVA, A. E.; GATTULLO, M.; DEBERNARDIS, S.; MONNO, G. Augmented reality on large screen for interactive maintenance instructions. **Computers in Industry**, Elsevier B.V., v. 65, n. 2, p. 270–278, fev. 2014. DOI: 10.1016/j.compind.2013.11.004.

FRIEDRICH, W. ARVIKA-augmented reality for development, production and service. **Proceedings. International Symposium on Mixed and Augmented Reality**. p. 3–4, 2002. DOI: 10.1109/ISMAR.2002.1115059.

GOLOMB, S. W. **Polyominoes: puzzles, patterns, problems, and packings**. 2. ed. [S.l.]: Princeton University Press, 1996. 184 p.

GOTTIPOLU, R. B.; GHOSH, K. A simplified and efficient representation for evaluation and selection of assembly sequences. **Computers in Industry**, v. 50, n. 3, p. 251–264, 2003. DOI: 10.1016/S0166-3615(03)00015-0.

HAKKARAINEN, M.; WOODWARD, C.; BILLINGHURST, M. Augmented assembly using a mobile phone. **2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality**, leee, p. 167–168, set. 2008. DOI: 10.1109/ISMAR.2008.4637349.

HENDERSON, S.; FEINER, S. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. **IEEE Transactions on Visualization and Computer Graphics**, leee, v. 17, n. 10, p. 1355–1368, 2011.

HOU, L.; WANG, X.; TRUIJENS, M. Using Augmented Reality to Facilitate Piping Assembly: An Experiment-Based Evaluation. **Journal of Computing in Civil Engineering**, p. 05014007, jul. 2013. DOI: 10.1061/(ASCE)CP.1943-5487.0000344.

KITAGAWA, M.; YAMAMOTO, T. 3D puzzle guidance in augmented reality environment using a 3D desk surface projection. **IEEE Symposium on 3D User Interfaces 2011**. p. 133–134, mar. 2011. DOI: 10.1109/3DUI.2011.5759241.

KREVELEN, D. W. F.; POELMAN, R. A Survey of Augmented Reality Technologies , Applications and Limitations. **The International Journal of Virtual Reality**, v. 9, n. 2, p. 1–20, 2010.

LI, Y.; KAMEDA, Y.; OHTA, Y. AR replay in a small workspace. **2013 23rd International Conference on Artificial Reality and Telexistence (ICAT)**. p. 97–101, dez. 2013. DOI: 10.1109/ICAT.2013.6728913.

MARQUARDT, D. W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. **Journal of the Society for Industrial and Applied Mathematics**, v. 11, n. 2, p. 431–441, jun. 1963. DOI: 10.1137/0111030.

MEBIUS, J. E. Derivation of the Euler-Rodrigues formula for three-dimensional rotations from the general formula for four-dimensional rotations. p. 1–7, jan. 2007.

MILGRAM, P.; KISHINO, F. A taxonomy of mixed reality visual displays. **IEICE Transactions on Information Systems**, E77-D, n. 12, p. 1–15, 1994.

MURAKAMI, K.; KIYAMA, R.; NARUMI, T.; TANIKAWA, T.; HIROSE, M. Poster: A wearable augmented reality system with haptic feedback and its performance in virtual assembly tasks. **2013 IEEE Symposium on 3D User Interfaces (3DUI)**. p. 161–162, mar. 2013. DOI: 10.1109/3DUI.2013.6550228.

NAKANISHI, M.; OZEKI, M.; AKASAKA, T.; OKADA, Y. Human factor requirements for Applying Augmented reality to manuals in actual work situations. **2007 IEEE International Conference on Systems, Man and Cybernetics**, IEEE, p. 2650–2655, 2007. DOI: 10.1109/ICSMC.2007.4413588.

NISHIHARA, A.; OKAMOTO, J. Object recognition in assembly assisted by augmented reality system. **SAI Intelligent Systems Conference (IntelliSys)**, 2015. p. 400–407, Nov 2015. DOI: 10.1109/IntelliSys.2015.7361172.

NOVAK-MARCINCIN, J.; BARNA, J.; JANAK, M.; NOVAKOVA-MARCINCINOVA, L.; TOROK, J. Visualization of intelligent assembling process by augmented reality tools application. **2012 4th IEEE International Symposium on Logistics and Industrial Informatics**. p. 33–36, set. 2012. DOI: 10.1109/LINDI.2012.6319505.

OKAMOTO, J.; NISHIHARA, A. Assembly assisted by augmented reality (a3r). In: **INTELLIGENT Systems and Applications: Extended and Selected Results from the SAI Intelligent Systems Conference (IntelliSys) 2015**. Cham: Springer International Publishing, 2016. p. 281–300. DOI: 10.1007/978-3-319-33386-1_14.

ONG, S.; WANG, Z. Augmented assembly technologies based on 3D bare-hand interaction. **CIRP Annals - Manufacturing Technology**, v. 60, n. 1, p. 1–4, jan. 2011. DOI: 10.1016/j.cirp.2011.03.001.

RACZYNSKI, A.; GUSSMANN, P. Services and training through augmented reality. **Visual Media Production, 2004.(CVMP). 1st European Conference on**. p. 263–271, 2004.

REINERS, D.; STRICKER, D.; KLINKER, G.; MÜLLER, S. Augmented reality for construction tasks: Doorlock assembly. **Proc. IEEE and ACM IWAR**, 1998.

SÄÄSKI, J.; SALONEN, T.; HAKKARAINEN, M.; SILTANEN, S.; WOODWARD, C.; LEMPÄINEN, J. Integration of Design and Assembly Using Augmented Reality. **Micro-Assembly Technologies and Applications**., v. 260, p. 395–404, 2008. DOI: 10.1007/978-0-387-77405-3_39.

SALONEN, T.; SÄÄSKI, J.; HAKKARAINEN, M.; KANNETIS, T.; PERAKAKIS, M.; SILTANEN, S.; POTAMIANOS, A.; KORKALO, O.; WOODWARD, C. Demonstration of assembly work using augmented reality. **Proceedings of the 6th ACM international conference on Image and video retrieval - CIVR '07**. p. 120–123, 2007. DOI: 10.1145/1282280.1282301.

SERVÁN, J.; MAS, F.; MENÉNDEZ, J.; RÍOS, J. Assembly Work Instruction Deployment Using Augmented Reality. **Key Engineering Materials**, v. 502, p. 25–30, fev. 2012. DOI: 10.4028/www.scientific.net/KEM.502.25.

SUZUKI, S.; ABE, K. Topological structural analysis of digitized binary images by border following. **Computer Vision, Graphics, and Image Processing**, v. 30, n. 1, p. 32–46, abr. 1985. DOI: 10.1016/0734-189X(85)90016-7.

TANG, A.; OWEN, C.; BIOCCA, F.; MOU, W. Comparative effectiveness of augmented reality in object assembly. **Proceedings of the conference on Human factors in computing systems - CHI '03**. p. 73, 2003. DOI: 10.1145/642611.642626.

WANG, Z.; SHEN, Y.; ONG, S. K.; NEE, A. Y.-C. Assembly Design and Evaluation Based on Bare-Hand Interaction in an Augmented Reality Environment. **2009 International Conference on CyberWorlds**. p. 21–28, 2009. DOI: 10.1109/CW.2009.15.

WANG, Z. B.; ONG, S. K.; NEE, a. Y. C. Augmented reality aided interactive manual assembly design. **The International Journal of Advanced Manufacturing Technology**, v. 69, n. 5-8, p. 1311–1321, jun. 2013. DOI: 10.1007/s00170-013-5091-x.

WESTERFIELD, G. **Intelligent Augmented Reality Training for Assembly and Maintenance**. 2012. 124 p. Tese (Master of Science) — University of Canterbury.

WIEDENMAIER, S.; OEHME, O.; SCHMIDT, L.; LUCZAK, H. Augmented reality (AR) for assembly processes design and experimental evaluation. **International Journal of Human-Computer Interaction**, v. 16, n. 3, p. 497–514, 2003.

YUAN, M. L.; ONG, S. K.; NEE, A. Y. C. Augmented reality for assembly guidance using a virtual interactive tool. **International Journal of Production Research**, v. 46, n. 7, p. 1745–1767, abr. 2008. DOI: 10.1080/00207540600972935.

ZAUNER, J.; HALLER, M.; BRANDL, A.; HARTMAN, W. Authoring of a mixed reality assembly instructor for hierarchical structures. **The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings**. p. 237–246, 2003. DOI: 10.1109/ISMAR.2003.1240707.

ZHANG, Z. Flexible camera calibration by viewing a plane from unknown orientations. **Proceedings of the Seventh IEEE International Conference on Computer Vision**. v. 1, p. 666–673, 1999. DOI: 10.1109/ICCV.1999.791289.