Vinicius Ferraz Guimarães

Estabilização de Imagens de Videolaringoscopia

Dissertação de mestrado apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia - Escola de Engenharia de São Carlos / Faculdade de Medicina de Ribeirão Preto / Instituto de Química de São Carlos da Universidade de São Paulo como parte dos requisitos para obtenção do título de mestre em Bioengenharia.

Área de Concentração: Bioengenharia

Orientador: Prof. Dr. José Carlos Pereira

São Carlos

2008

*”São fúteis e cheias de erros as ciências que não nasceram da experimentação, mãe de
todo conhecimento.”*

Leonardo da Vinci

Agradecimentos

À meus pais pelo apoio e incentivo.

Ao professor Dr. José Carlos Pereira pela orientação.

Ao professor Dr. Arlindo Neto Montagnoli pela ajuda nos algoritmos.

À meus amigos do LIM Maria Eugênia, Daiane, Rafael, Jullian, Soledad, Edwin, Paulo e Everthon.

À CAPES pelo apoio financeiro.

À todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

Resumo

GUIMARÃES, V. F., **Estabilização de Imagens de Videolaringoscopia**. 2008. 48f. Programa de Pós-Graduação Interunidades em Bioengenharia - EESC/FMRP/IQSC da Universidade de São Paulo, São Carlos, 2008.

Nesse trabalho foram utilizadas e analisadas diferentes técnicas computacionais para estabilização de imagens de videolaringoscopia, adquiridas por meio de videostroboscopia, a fim de remover os movimentos indesejados provenientes do operador da câmera e do paciente durante a gravação, de modo a melhorar sua qualidade. Com um vídeo estabilizado podemos obter uma quimografia, que é um exame que permite observar os padrões de abertura e fechamento das pregas vocais durante a fonação, com boa qualidade. As técnicas utilizadas na estabilização foram correlação espacial, correlação de fase e *Optical Flow*.

Palavras-Chave: Estabilização de imagens, videolaringoscopia, quimografia.

Abstract

GUIMARÃES, V. F., **Image Stabilization for Videolaryngoscopy**. 2008. 48f. Programa de Pós-Graduação Interunidades em Bioengenharia - EESC/FMRP/IQSC da Universidade de São Paulo, São Carlos, 2008.

In this work, different computational techniques were used and evaluated for image stabilization of videolaryngoscopy scenes, so that the unwanted movements coming from cameraman and patient's movements can be removed, thus improving their quality. After stabilizing the video, we can get a kymography, an examination which allows us to visualize the vibratory pattern of vocal fold, with a good quality. The techniques used were spatial correlation, phase correlation and Optical Flow.

Key-words: Image stabilization, videolaryngoscopy, kymography.

Lista de Figuras

1.1	Estruturas do aparelho fonador.	13
1.2	Estruturas da laringe.	14
1.3	Pregas vocais e glote.	14
1.4	Videolaringoscopia	15
1.5	A imagem de vídeo, a linha selecionada para análise e sua respectiva quimografia projetada abaixo.	16
1.6	Parâmetros extraídos da quimografia.	16
1.7	Princípio de funcionamento da videoestroboscopia.	17
2.1	Sistema de estabilização de imagem mecânico.	20
2.2	Sistema STEADICAM®.	20
2.3	Sistema de estabilização de imagem óptico.	21
2.4	Direção do gradiente do brilho.	23
2.5	Problema da abertura.	23
2.6	As três derivadas parciais da imagem.	24
2.7	Processo de correlação e template.	26
2.8	Matriz com os valores da correlação.	26
2.9	Superfície de correlação de fase.	28
3.1	Fluxograma do método de correlação espacial.	30
3.2	Região de escolha do <i>template</i>	31
3.3	Fluxograma do método de correlação de fase.	31
3.4	Localização das sub-imagens.	32
3.5	Vetor de deslocamento local $\vec{d}$	33
3.6	<i>Frame</i> do vídeo com suas respectivas sub-imagens.	33
3.7	Fluxograma do método do <i>Optical Flow</i>	34

3.8	Sub-regiões para determinação do <i>Optical Flow</i>	35
4.1	Movimento antes e após a estabilização no eixo X.	37
4.2	Movimento antes e após a estabilização no eixo Y.	37
4.3	Quimografia realizada antes e após da estabilização do vídeo.	37
4.4	Deslocamento dos frames na direção x antes e depois da estabilização. . . .	38
4.5	Deslocamento dos frames na direção y antes e depois da estabilização. . . .	39
4.6	Quimografia realizada antes e após da estabilização do vídeo.	39
4.7	Movimento antes e após a estabilização no eixo X utilizando <i>Optical Flow</i> . .	40
4.8	Movimento antes e após a estabilização no eixo Y utilizando <i>Optical Flow</i> . .	40
4.9	Quimografia realizada antes e após da estabilização do vídeo.	41

Lista de Tabelas

1.1	Comparação entre os métodos de vídeo de alta velocidade, videoestroboscopia e videoquimografia.	18
4.1	Variância dos deslocamentos antes a depois da estabilização utilizando a técnica de correlação espacial.	36
4.2	Variância dos deslocamentos antes a depois da estabilização utilizando a técnica de correlação de fase.	38
4.3	Variância dos deslocamentos antes a depois da estabilização utilizando a técnica de <i>Optical Flow</i>	39
4.4	Comparação da variância entre os três métodos CE - Correlação espacial, CF - Correlação de Fase e OF - <i>Optical Flow</i>	41

Conteúdo

1	Introdução	11
1.1	Estruturas do aparelho fonador	12
1.2	Videolaringoscopia	15
1.3	Quimografia	16
1.4	Objetivos	18
2	Estabilização de Imagens	19
2.1	Sistemas Mecânicos	19
2.2	Sistema Óptico	20
2.3	Sistemas Computadorizados	21
2.3.1	Optical Flow	21
2.3.2	Features Tracking	24
2.3.3	Correlação Espacial	25
2.3.4	Correlação de Fase	25
3	Materiais e Métodos	29
3.1	Correlação Espacial	29
3.2	Correlação de Fase	31
3.3	<i>Optical Flow</i>	33
4	Resultados	36
4.1	Correlação Espacial	36
4.2	Correlação de fase	38
4.3	<i>Optical Flow</i>	39
4.4	Discussão	41

5	Conclusões	43
5.1	Trabalhos futuros	44
	Referências	45

Capítulo 1

Introdução

A emissão de voz é um fenômeno muito comum que faz parte do nosso dia-a-dia, mas entender o seu funcionamento é uma tarefa ainda muito complexa (Koishi et al., 2003).

A produção de voz requer uma unidade funcional obtida por meio da ação integrada de uma série de estruturas que não participam da mesma unidade orgânica, pelo fato de não existir um órgão especialmente desenvolvido para a produção da fala (Montagnoli, 2006).

Dentre essas estruturas, as pregas vocais são aquelas que apresentam um mecanismo de funcionamento mais complexo e delicado. Devido as interações de mecanismos não-lineares encontrados nas pregas vocais, essas estruturas são capazes de assumir várias características biomecânicas e produzir uma grande variedade de sons (Stevens, 1977).

De acordo com (Ishizaka e Matsudaria, 1972) os parâmetros fonatórios como frequência, duração, timbre e intensidade podem sofrer variações de acordo com a alteração dos seguintes fatores:

1. pressão aérea na subglote;
2. fluxo de ar passando pela glote;
3. resistência glótica;
4. contração dos músculos das pregas vocais;
5. padrão de fechamento;
6. massa vibrátil, rigidez e elasticidade das pregas vocais;
7. acoplamento acústico das cavidades acima e abaixo das pregas vocais (caixas de ressonâncias);

Quando ocorre uma alteração no sistema de produção de voz, essa geralmente ocorre na laringe, ou especificamente nas pregas vocais; produzindo um dos sintomas mais freqüentes nas alterações vocais, a disfonia.

A disfonia é um sintoma decorrente de patologias como nódulos, pólipos, paralisias e edemas nas pregas vocais que contribuem com as alterações dos parâmetros fonatórios (Kim et al., 2003), sendo que em, alguns casos, intervenções cirúrgicas são necessárias. Além das mencionadas patologias, próprias do sistema vocal, fatores funcionais ou fisiológicos, hereditários, genéticos, ambientais, psicológicos, entre outros podem interferir no correto funcionamento do complexo sistema de produção de voz.

O diagnóstico precoce das alterações laríngeas é fundamental para o tratamento e prognóstico de pacientes com problemas vocais. Atualmente existem variadas técnicas e ferramentas diagnósticas empregadas na clínica vocal; uma das mais modernas é o exame de quimografia, que permite observar a variabilidade da vibração das pregas vocais através de imagens de vídeo.

1.1 Estruturas do aparelho fonador

Os sons da fala humana são resultado da ação do sistema de órgãos que constituem o aparelho fonador, formado por pulmões, traquéia, laringe, pregas vocais, faringe, cavidade bucal e nasal (caixas de ressonância) e lábios, por onde o som é irradiado. As estruturas constituintes do aparelho fonador são mostradas na figura 1.1, com exceção dos pulmões e pregas vocais.

1. Cavidade nasal
2. Palato duro
3. Lábio superior
4. Dente incisivo
5. Lábio inferior
6. Mandíbula
7. Língua

Figura 1.1: Estruturas do aparelho fonador.

- 8. Osso da Hióide
- 9. Nasofaringe
- 10. Palato mole
- 11. Úvula
- 12. Faringe
- 13. Epiglote
- 14. Laringe
- 15. Esôfago

Os pulmões são responsáveis por gerar a corrente de ar que penetra pela traquéia e chega a laringe. A laringe, localizada na região ântero-superior do pescoço, adiante da faringe e acima da traquéia, é exibida na figura 1.2. Nela fica localizada a glote, exibida na figura 1.3, que é o espaço situado entre as duas pregas vocais; a epiglote cuja função é proteger os pulmões da entrada de líquidos e sólidos. Ao atravessar a glote, sob condições de estiramento das pregas vocais, o fluxo de ar causa vibrações das mesmas, criando, após ressonância do trato vocal, os sons vocálicos. Sob condições de relaxamento das pregas,

a passagem de ar pela glote não produz vibrações das pregas vocais e, após equalização no trato vocal, resulta em articulações surdas.

Ao passar pela laringe o ar chega a faringe que é responsável por sua condução para o canal bucal ou nasal. O palato mole, quando o mesmo levanta-se e cola a parte posterior da faringe, é responsável pela produção dos sons não-nasais, deixando apenas a cavidade bucal livre para a passagem do ar. Quando ele é abaixado, as fossas nasais são conectados ao sistema fonatório e, dessa maneira, produz os sons nasalados. Por fim temos a boca e o nariz que tem a função de irradiar os sons produzidos para o ambiente.

Figura 1.2: Estruturas da laringe. Fonte (*Anatomy & Physiology of the Larynx*, Acessado 1 junho de 2007).

Figura 1.3: Pregas vocais e glote.

1.2 Videolaringoscopia

Durante a fonação as pregas vocais desempenham uma importante função, que é controlar a passagem de ar através da laringe, de modo que elas trabalhem como um oscilador que interage com a pressão do ar no trato respiratório (Lee. et al., 2001).

A vibração das pregas vocais determina a qualidade da voz, dessa maneira, o objetivo de avaliá-la e quantificá-la é uma importante tarefa. A observação laringoscópica das pregas vocais oferece o mais eficiente método para a investigação do comportamento das mesmas (Švec, 2000).

Pelo fato das pregas vocais e a laringe situarem-se em uma região de difícil acesso, técnicas especiais são necessárias para examinar suas estruturas e funcionamento.

Dentre elas podemos citar a videolaringoscopia, que consiste em uma câmera de vídeo acoplado a um laringoscópio de modo que imagens da laringe e das pregas vocais possam ser capturadas para uma posterior análise. A figura 1.4 mostra como o exame de videolaringoscopia é realizado. Um endoscópio rígido, cuja extremidade possui uma câmera e uma fonte de luz, é introduzida na boca do paciente de modo possibilitar a filmagem das pregas vocais. As imagens são gravadas em um videocassete, ou capturadas a partir de uma placa de vídeo de um microcomputador, de modo que possam ser processadas posteriormente.

A partir dessa técnica é realizado um exame denominado quimografia, cujo objetivo é observar os padrões de abertura e fechamento das pregas vocais durante a fonação.

Figura 1.4: Videolaringoscopia. Fonte (Instruments, Acessado em 1 Julho de 2007).

1.3 Quimografia

O exame de quimografia nos permite fazer uma análise quantitativa da vibração das pregas vocais (Sung et al., 1999). Ele consiste na junção de uma sequência de linhas obtidas a partir dos frames do vídeo, capturados através da videolaringoscopia, figura 1.5. Para isso devemos selecionar uma linha de análise na região onde apresenta uma maior amplitude de abertura das pregas vocais.

Figura 1.5: A imagem de vídeo, a linha selecionada para análise e sua respectiva quimografia projetada abaixo.

A partir deste exame podem ser feitas medidas como dos tempos de fase aberta (TFA), fase fechada (TFF), abertura (T_a), fechamento (T_f) e amplitude de abertura (A), conforme mostrado na figura 1.6, PVD e PVE pregas vocais direita e esquerda respectivamente.

Figura 1.6: Parâmetros extraídos da quimografia. Fonte (Koishi et al., 2003).

O exame de quimografia inicia-se filmando a vibração das pregas vocais através da videolaringoscopia. Existem três variações do modo que a captura de vídeo pode ser feita, o que vai diferir entre elas é o tipo de câmera utilizado para fazê-la, que pode ser a câmera de alta velocidade, videoestroboscopia e videoquimógrafo.

A câmera de alta velocidade, capaz de capturar entre 1000 a 4000 frames por segundo aproximadamente, registra o verdadeiro movimento das pregas vocais durante cada ciclo de vibração além de exibir a imagem das pregas vocais como um todo.

A videoestroboscopia, que utiliza uma câmera comum, capturando entre 25 e 30 quadros por segundo, fica acoplada a uma fonte de luz estroboscópica. O estrobo produz flashes de luz com uma frequência submúltipla da vibração das pregas vocais, ocasionando um movimento ilusório de câmera lenta, e dessa forma é possível observar as mudanças nos movimentos da mucosa das pregas vocais (Kim et al., 2003). A figura 1.7 ilustra o princípio de funcionamento da videoestroboscopia.

Entretanto esse método apresenta uma limitação, ele pode somente registrar vibrações periódicas. Dessa maneira, vibrações irregulares não poderão ser avaliadas corretamente (Švec e Schutte, 1996).

Figura 1.7: Princípio de funcionamento da videoestroboscopia. Os flashes de luz são disparados uma vez em cada frame do vídeo em um determinado momento (acima). A imagem capturada a partir de cada frame são gravadas para criar um ciclo artificial (abaixo).

O método por videoquimógrafo é capaz de registrar até 8000 linhas por segundo, ou seja, o videoquimógrafo filma somente uma linha do vídeo de cada vez.

A tabela 1.1 mostra as principais características de cada método de videolaringoscopia.

Tabela 1.1: Comparação entre os métodos de vídeo de alta velocidade, videoestroboscopia e videoquimografia. Fonte (Hertegard et al., 2003)

	Câmera de alta velocidade	Videoestroboscopia	Videoquimografia
Resolução temporal (imagens por segundo)	1000-8000	25-30	8000
Tamanho da imagem (pixel x pixel)	256x256	700x500	700x1
Cor	Não	Sim	Não
Preço aproximado(U\$)	18.000,00	13.000,00	8.000,00
Armazenamento (tempo aproximado)	Digital, 256Mb 2 seg.	Analógica 5 min.	Analógica 5 min.

1.4 Objetivos

Os objetivos deste trabalho são:

- Avaliar sistemas de estabilização de imagens aplicado a videolaringoscopia.
- Melhorar as imagens para permitir exames de quimografias mais precisos.

Capítulo 2

Estabilização de Imagens

Os sistemas de estabilização de imagens são utilizados para remover movimentos indesejados nas seqüências de vídeo, isto é, movimentos de rotação, translação e zoom (escala), aplicando algum tipo de transformação geométrica nas imagens. O tipo mais comum de degradação é causado pelo movimento de translação, sendo que a maioria dos sistemas de estabilização tenta eliminar somente esse movimento (Yeni e Erturk, 2005).

Os sistemas de estimação podem ser classificados de três maneiras: sistemas mecânicos, ópticos ou computadorizados.

2.1 Sistemas Mecânicos

Os sistemas mais antigos baseavam-se em sistemas mecânicos, onde sensores giroscópicos detectam a velocidade e a direção do movimento e atuadores movem o conjunto de lentes e CCD. A figura 2.1 mostra o sistema mecânico proposto por (Oshima et al., 1989).

Outro sistema é o chamado STEADICAM[®], que consiste em um suporte no qual a câmera é fixada e presa ao corpo do operador com a ajuda de um colete. Tal suporte apresenta uma grande massa inercial que não pode ser facilmente movida por pequenos movimentos (como por exemplo no momento de ajustar a câmera) e nem movimentada bruscamente, dessa maneira permitindo apenas os movimentos suaves, esse sistema é mostrado na figura 2.2.

Figura 2.1: Sistema de estabilização de imagem mecânico.

Figura 2.2: Sistema STEADICAM®.

2.2 Sistema Óptico

No método óptico (Sato et al., 1993) em que o sistema de detecção dos movimentos é igual ao do sistema mecânico porém a compensação do movimento é feita através da deformação da lente objetiva da câmera, na qual é constituída por um par de placas transparentes de vidro acopladas por um suporte flexível preenchido por líquido. Cada placa de vidro é movimentada de forma independente por atuadores na direção vertical, horizontal ou rotacionada. Dessa forma, fazendo com que a luz proveniente do objeto seja desviada em sentido oposto ao movimento da câmera, a imagem é estabilizada. A figura 2.3 mostra o conjunto de lentes desse sistema desligado e ativado, respectivamente.

Figura 2.3: Sistema de estabilização de imagem óptico.

2.3 Sistemas Computadorizados

Os modernos sistemas de estabilização de imagens são baseados em métodos computacionais para processamento de imagens. Estes são compostos basicamente por duas etapas: a estimação do movimento e correção.

Para a estimação do movimento existem várias técnicas, dentre elas podemos citar *Optical Flow*, *Features Tracking*, correlação espacial e correlação de fase.

Para correção é utilizado a transformação *affine* de translação, para deslocar a imagem no sentido oposto daquela estimada, representada conforme equação 2.1 (Gonzalez e Woods, 1993):

$$\begin{bmatrix} X' \\ Y' \end{bmatrix} = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (2.1)$$

onde X e Y são as coordenadas dos pontos da imagem original, X' e Y' são as coordenadas dos pontos da imagem transladada e dx e dy são a quantidade de pixels que serão deslocados no eixo x e y , respectivamente.

2.3.1 Optical Flow

Optical Flow ou *Image Flow* é a distribuição da velocidade, em relação ao observador, sobre os pontos de uma imagem. Isso nos dá informações úteis para análise da dinâmica

da cena (Jain et al., 1995). Se determinarmos o *Optical Flow* a partir de duas imagens consecutivas, obtemos um Vetor de Deslocamento (VD) entre elas (Jahne, 1997).

A relação entre *Optical Flow* no plano da imagem e as velocidades dos objetos no mundo tridimensional nem sempre existe (Horn e Schunk, 1981). Por exemplo, considere uma esfera com uma superfície e iluminação uniforme. Se ela for rotacionada a imagem não apresentará *Optical Flow* em nenhum ponto, a menos que a iluminação mude. Considere o caso contrário em que a esfera não esta rotacionada mas a iluminação esta sofrendo alguma mudança. A imagem portanto apresentará *Optical Flow* em algum ponto dela. Sendo assim superfícies que apresentem pouco contraste ou uniformes apresentarão problemas na estimação de movimento utilizando *Optical Flow*.

Muitas formas de se calcular *Optical Flow* a partir de duas ou mais imagens foram desenvolvidas. Dentre elas podemos citar o Método do Gradiente.

O Método do Gradiente explora a relação entre o gradiente temporal e espacial da intensidade da imagem.

Suponha que a intensidade de um ponto (x,y) no plano da imagem é dado como $E(x,y,t)$. Assumindo pequenos movimentos, a intensidade nesse ponto irá permanecer constante, então (Jain et al., 1995):

$$\frac{dE}{dt} = 0 \quad (2.2)$$

Usando a regra da cadeia para diferenciação temos que:

$$\frac{\partial E}{\partial x} \frac{dx}{dt} + \frac{\partial E}{\partial y} \frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \quad (2.3)$$

usando

$$u = \frac{dx}{dt} \quad (2.4)$$

$$v = \frac{dy}{dt} \quad (2.5)$$

a relação entre o gradiente espacial e temporal e as componentes de velocidade é:

$$E_x u + E_y v + E_t = 0 \quad (2.6)$$

onde E_x , E_y e E_t são as derivadas parciais do brilho da imagem em função de x , y e t respectivamente, e podem ser calculadas diretamente a partir da imagem. Dessa maneira

há duas variáveis desconhecidas, u e v , e somente uma equação. Sendo assim o *Optical Flow* não pode ser determinado usando somente as informações contidas no ponto.

As velocidades (u,v) estão na direção perpendicular ao vetor do gradiente do brilho, (E_x, E_y) , e seu módulo, $\Theta(E_x, E_y)$, é igual a E_t dividido pela magnitude de (E_x, E_y) (Horn e Schunk, 1981).

$$\Theta(E_x, E_y) = -\frac{E_t}{\sqrt{E_x^2 + E_y^2}} \quad (2.7)$$

Dessa maneira não podemos determinar o movimento na direção da linha de brilho constante, mostrado na figura 2.4, assim podemos estimar apenas a componente normal a esta linha, esse fenômeno é conhecido como *problema de abertura*, em que podemos detectar o movimento mas não suas componentes de velocidades u e v .

Figura 2.4: Direção do gradiente do brilho.

Figura 2.5: Problema da abertura, se observarmos uma área muito pequena o movimento dessa não pode ser determinada corretamente.

Para estimar as derivadas parciais $\frac{\partial E}{\partial x}$, $\frac{\partial E}{\partial y}$ e $\frac{\partial E}{\partial t}$, denotadas por E_x , E_y e E_t respecti-

vamente, usamos o método utilizado por (Horn e Schunk, 1981), onde $E_{i,j,k}$ é o brilho na i -ésima linha, j -ésima coluna e k -ésimo frame.

$$E_x = \frac{\{E_{i,j+1,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i+1,j,k} + E_{i,j+1,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i+1,j,k+1}\}}{4} \quad (2.8)$$

$$E_y = \frac{\{E_{i+1,j,k} - E_{i,j,k} + E_{i+1,j+1,k} - E_{i,j+1,k} + E_{i+1,j,k+1} - E_{i,j,k+1} + E_{i+1,j+1,k+1} - E_{i,j+1,k+1}\}}{4} \quad (2.9)$$

$$E_t = \frac{\{E_{i,j,k+1} - E_{i,j,k} + E_{i+1,j,k+1} - E_{i+1,j,k} + E_{i,j+1,k+1} - E_{i,j+1,k} + E_{i+1,j+1,k+1} - E_{i+1,j+1,k}\}}{4} \quad (2.10)$$

Figura 2.6: As três derivadas parciais da imagem no centro do cubo são cada uma estimada a partir da média da diferença ao longo das quatro bordas paralelas do cubo.

2.3.2 Features Tracking

Há a técnica de *Feature Tracking* (Censi et al., 1999), no qual uma série de pontos característicos (pontos, linhas ou cantos de controle) são selecionados em uma imagem, geralmente estes pontos são locais que apresentam um grande contraste e, comparados estabelecendo uma correspondência com uma outra série de pontos de outra imagem dessa

maneira determinando o deslocamento entre elas. Entretanto a escolha dos pontos e o estabelecimento dessas correspondências não é uma tarefa fácil (Jain et al., 1995). Esta técnica costuma ser utilizada para construção de mapas a partir de imagens de satélites (Fedorov, 2002).

2.3.3 Correlação Espacial

Para fazer a correlação espacial, foi selecionado um frame de referência onde um *template* é comparada com toda a imagem, *frame a frame*. Para cada uma dessas interações é obtida uma matriz de correlação. Cada uma dessas matrizes tem um pico de máxima amplitude, então o processo de alinhamento consiste em deslocar todos esses picos para a região de máximo de um frame de referência. A correlação é obtida pela seguinte equação:

$$R(x, y) = \frac{\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x', y') I(x + x', y + y')}{\left[\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x', y')^2 \sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} I(x + x', y + y')^2 \right]^{\frac{1}{2}}} \quad (2.11)$$

onde:

$I(x, y)$ é o valor do pixel da imagem na posição (x, y) ;

$T(x, y)$ é o valor do pixel do template na posição (x, y) ;

h e w são a altura e a largura do template respectivamente.

A figura 2.7 mostra como é o processo do cálculo da correlação e figura 2.8 mostra a matriz de correlação, onde o ponto branco indicado pela seta corresponde ao valor máximo.

2.3.4 Correlação de Fase

As três técnicas citadas anteriormente trabalham no domínio espacial, há também as que trabalham no domínio da frequência, tais como a técnica de correlação de fase mostrada pela primeira vez por (Kuglin e Hines, 1975) e implementada posteriormente por (Pearson et al., 1977; de Castro e Morandi, 1987; Reddy e Chatterji, 1996; Erturk e Dennis, 2000). Dentre suas vantagens destacam-se: detecção de grandes deslocamen-

Figura 2.7: Processo de correlação e template.

Figura 2.8: Matriz com os valores da correlação.

tos com grande acurácia, insensibilidade a ruídos e imunidade a mudanças de iluminação (Erturk e Dennis, 2000).

Esta técnica se baseia no fato de que a maior parte da informação a respeito do deslocamento relativo entre duas imagens está contida na fase dos seus espectros de correlação cruzada (Pearson et al., 1977).

A correlação de fase é baseada na propriedade do deslocamento da transformada de Fourier, na qual um deslocamento no domínio espacial causa um deslocamento linear na fase (domínio da frequência). Considere duas imagens $f(x,y)$ e $g(x,y)$ sendo esta última uma versão deslocada da primeira em (x_0, y_0) , isto é:

$$g(x, y) = f(x - x_0, y - y_0) \quad (2.12)$$

Assim a transformada de Fourier de $g(x,y)$ será:

$$G(\xi, \eta) = e^{-i2\pi(\xi x_0 + \eta y_0)} F(\xi, \eta) \quad (2.13)$$

onde $F(\xi, \eta)$ e $G(\xi, \eta)$ são as transformadas de Fourier das imagens $f(x, y)$ e $g(x, y)$ respectivamente e os índices ξ e η são as frequências espaciais.

A função de correlação entre as imagens $f(x, y)$ e $g(x, y)$ é dada por:

$$C(u, v) = \sum_{u=0}^{U-1} \sum_{v=0}^{V-1} f(x, y) g^*(x + u, y + v) \quad (2.14)$$

De acordo com o teorema da correlação podemos reescrever a equação 2.14 da seguinte maneira:

$$C(u, v) = \mathcal{F}^{-1}[F(\xi, \eta) G^*(\xi, \eta)] \quad (2.15)$$

onde $\mathcal{F}^{-1}[\cdot]$ representa o operador da transformada inversa de Fourier e $*$ denota o conjugado complexo da função.

As transformadas das imagens são funções complexas e podem ser escritas do como:

$$F(\xi, \eta) = |F(\xi, \eta)| e^{i\phi_F(\xi, \eta)} \quad (2.16)$$

$$G(\xi, \eta) = |G(\xi, \eta)| e^{i\phi_G(\xi, \eta)} \quad (2.17)$$

A partir das equações 2.16 e 2.17 podemos obter a diferença relativa de fase entre as duas imagens do seguinte modo (Ohya et al., 1987) :

$$e^{i(\phi_F(\xi, \eta) - \phi_G(\xi, \eta))} = \frac{F(\xi, \eta) G^*(\xi, \eta)}{|F(\xi, \eta) G^*(\xi, \eta)|} \quad (2.18)$$

Para se obter vetor de deslocamento da imagem fazemos a transformada de Fourier inversa da equação 2.18. Desta forma obtemos uma *superfície de correlação de fase*, que possui um pico na região correspondente ao vetor deslocamento entre as imagens, ou seja, uma função impulso deslocada (dx, dy) unidades.

$$\mathcal{F}^{-1}[e^{i(\phi_F(\xi, \eta) - \phi_G(\xi, \eta))}] = \delta(x - dx, y - dy) \quad (2.19)$$

A figura 2.9 mostra a superfície de correlação de fase das imagens $f(x, y)$ e $g(x, y)$.

Figura 2.9: Superfície de correlação de fase.

Capítulo 3

Materiais e Métodos

Os vídeo utilizados no trabalho foram fornecidos pelo ambulatório de Otorrinolaringologia do Hospital das Clínicas de São Paulo (HC-FMSP), tem o tamanho de 320x240 pixels, no formato avi. Foram utilizados vídeos de aproximadamente 4 segundos cada.

O sistema foi desenvolvido em um microcomputador sob o sistema operacional Windows XP, o compilador utilizado foi o Borland C++ Builder 4.0 com a biblioteca *OpenCV* da Intel(“Open Source Computer Vision Library”) (*Open Source Computer Vision Library.*, Acessado em 10 de março de 2007) dedicada ao processamento de imagens.

Nesse trabalho foram testadas três técnicas para fazer a estabilização de imagens: Correlação espacial, correlação de fase e *Optical Flow*.

3.1 Correlação Espacial

Para fazer a estabilização de imagens utilizando esta técnica, primeiramente o vídeo é convertido para preto e branco, dessa forma reduzimos o número de canais de cor de três para somente um, que é o suficiente para conter as informações a sobre o movimento das imagens. Após a conversão para cinza foi feito uma equalização de histograma de tons de cinza para aumentar o contraste e melhorar a definição das estruturas na imagem, em seguida foi escolhido um *frame* de referência, este é extraído no momento em que a luz estroboscópica começa a funcionar.

Após escolher o *frame* de referência um pequeno *template* foi extraído, com tamanho de 46x46 pixels escolhida empiricamente, a região onde o mesmo é extraído deve ser selecionado de um lugar que possua um bom contraste entre as partes, geralmente localizada

na região posterior da laringe, conforme mostrada na figura 3.2. Essa escolha empírica se deu pelo fato de não haver na literatura um método capaz de localizá-la e atribuir-lhe um tamanho, dessa maneira foram testadas regiões em diferentes posições e tamanhos, sendo que esses parâmetros podem variar de acordo com as dimensões do vídeo.

Em seguida é feita a correlação do *template* com um frame de referência utilizando a equação 2.11, com isso obtemos uma matriz de correlação. Encontrando seu ponto de máximo valor P_{0max} fazemos uso desse um ponto de referência para a determinação do movimento entres os *frames* seguintes.

Posteriormente, repetimos o processo de correlação do *template* com cada *frame* do vídeo, com isso obtemos outras matrizes de correlação nas quais deve-se localizar seus respectivos pontos de máximos P_{imax} .

O processo para determinar os deslocamentos entre os *frames* consiste em encontrar a diferença de localização entre P_{0max} e P_{imax} e corrigi-los aplicando a transformação *affine* da equação 2.1.

Figura 3.1: Fluxograma do método de correlação espacial.

Figura 3.2: Região de escolha do *template*.

3.2 Correlação de Fase

A técnica de correlação de fase foi utilizada para fazer a estabilização das imagens obtidas a partir das imagens das videoestroboscopia.

O método foi aplicado seguindo o seguinte esquema ilustrado na figura 3.3.

Figura 3.3: Fluxograma do método de correlação de fase.

Cada *frame* da seqüência do vídeo é convertido para preto e branco, foi feita a equalização de histograma, em seguida quatro sub-imagens são selecionadas, conforme mostra a figura 3.4. Estas sub-imagens são utilizadas para determinar o vetor de deslocamento local usando a correlação de fase. Para efetuar os cálculos das transformadas de Fourier (FFT), essas sub-imagens são quadradas de dimensões 64x64 pixels. Essa dimensão foi escolhida por apresentar baixa carga computacional e ser suficiente para estimar adequadamente os vetores de deslocamento (Erturk, 2003).

Figura 3.4: Localização das sub-imagens para determinação dos vetores de deslocamento local, extraído de (Erturk, 2003).

Para todas as sub-imagens de um *frame*, os vetores de deslocamento são encontrados a partir das respectivas sub-imagens em relação ao *frame* anterior.

Para cada superfície de correlação de fase, correspondendo as quatro sub-imagens, é encontrado o pico de maior amplitude P_n (onde n corresponde a sub-imagem) e em seguida esses valores são armazenados. Esse pico encontra-se em uma posição correspondente ao vetor deslocamento local. Para determinar o vetor de deslocamento global (VDG) do *frame*, figura 3.5, encontra-se o pico com maior amplitude e compara-se com os demais. Estes serão descartados se apresentarem amplitude menor que a metade do maior pico, ou seja, $P_n < \frac{\max(P_n)}{2}$, caso contrário será utilizado para calcular o VDG que será o valor médio dos valores de deslocamento locais que foram aceitos anteriormente.

A figura 3.6 mostra um frame de uma seqüência de vídeo capturado em um exame de videolaringoscopia. Quatro sub-imagens foram usadas para a determinação dos vetores de deslocamento local e o vetor de deslocamento global.

Foram feitas as correlações de fase em sub-imagem do *frame k* com a respectiva sub-imagem do frame *k-1*, obtendo-se quatro vetores de deslocamento local para cálculo pos-

Figura 3.5: Vetor de deslocamento local $\vec{d}$.

Figura 3.6: *Frame* do vídeo com suas respectivas sub-imagens.

terior do vetor de deslocamento global.

3.3 *Optical Flow*

Nessa etapa foi utilizado o método do *Optical Flow* de Horn e Schunck (Horn e Schunck, 1981) para fazer a estabilização das imagens, seguindo a abordagem proposta por (Vella et al., 2002). Esta abordagem é semelhante a utilizada na técnica de correlação de fase, onde são selecionadas sub-regiões da imagem para fazer o seu processamento.

O esquema de funcionamento segue os seguinte passos da figura 3.7.

Primeiramente as imagens são convertidas para preto e branco, foi feita a equalização de histograma e em seguida deve ser escolher as sub-regiões para ser processada. Foram escolhidas sub-regiões localizadas na parte superior e inferior dos *frames* conforme mostrado na figura 3.8, cada uma com um tamanho de 150x50 pixels, escolhida empiricamente. Em

seguida é calculado o *Optical Flow* entre elas e feito um histograma para saber qual a distribuição dos deslocamentos, sabemos que em áreas homogêneas nas imagens o *Optical Flow* pode ser nulo, assim temos que eliminar a componente do histograma que representa o deslocamento nulo, de modo a diminuir o erro de estimação de deslocamento. Em seguida temos que encontrar qual é a componente de velocidade que acontece com maior frequência, este será nosso vetor de deslocamento local. Repetindo o processo na outra sub-imagem, encontramos o segundo vetor de deslocamento local, com isso fazemos a média entre as duas e encontramos o vetor de deslocamento global do *frame*.

Figura 3.7: Fluxograma do método do *Optical Flow*.

Figura 3.8: Sub-regiões para determinação do *Optical Flow*.

Capítulo 4

Resultados

4.1 Correlação Espacial

Utilizando a técnica de correlação espacial obtivemos bons resultados, conforme mostrado nos gráficos das figuras 4.1 e 4.2, que mostra os movimentos antes e após a estabilização das imagens de vídeo. Nos gráficos podemos observar alguns picos na estimação do movimento que foram causados por causa de ruídos e grandes variações da iluminação.

Entretanto, o que mais dificulta a utilização desse método é escolha do tamanho e a localização do *template*, que deve conter áreas que apresente um alto contraste, devendo ser escolhida manualmente, dificultando a sua utilização.

As figuras 4.3 (a) e 4.3 (b) mostram os resultados das quimografias realizadas antes e após o processo de estabilização, podendo-se notar a melhora visual obtida.

A tabela 4.1 mostra o comparativo da variância do deslocamento antes e após a estabilização do vídeo.

Tabela 4.1: Variância dos deslocamentos antes e depois da estabilização utilizando a técnica de correlação espacial.

	X antes	X depois	Y antes	Y depois	Variação X (%)	Variação Y (%)
Variância	16.44	0.36	65.61	0.05	-97.83	-99.92

Figura 4.1: Movimento antes e após a estabilização no eixo X.

Figura 4.2: Movimento antes e após a estabilização no eixo Y.

Figura 4.3: Quimografia realizada antes (a) e após da estabilização do vídeo (b).

4.2 Correlação de fase

O resultado obtido para os deslocamentos acumulados nos eixos x e y estão indicados nas figuras (4.4) e (4.5), mostrando que o sistema foi capaz de eliminar parte dos movimentos de translação em ambos os eixos. A tabela 4.2 mostra a variância dos movimentos antes e depois da estabilização.

Tabela 4.2: Variância dos deslocamentos antes e depois da estabilização utilizando a técnica de correlação de fase.

	X antes	X depois	Y antes	Y depois	Varição X (%)	Varição Y (%)
Variância	6.82	11,56	11.18	8.40	69.50	-24,86

A pouca precisão da estimação do movimento pode ter ocorrido pelo fato de cada região apresentar diferente vetor de deslocamento local com direção e/ou amplitude distinta, em consequência do movimento não uniforme da imagem como um todo. O vetor médio calculado não representa o verdadeiro movimento global do *frame* do vídeo. Some-se ainda o fato de que as imagens do vídeo apresentam um baixo contraste, dificultando ainda mais a estimação do movimento.

A figura 4.6 (a) mostra a imagem de uma quimografia feita a partir de um vídeo antes de ser estabilizado, enquanto a figura 4.6 (b) mostra a quimografia do mesmo vídeo após o processo de estabilização. Comparando-se ambas pode se observar que a técnica utilizada corrige parcialmente a movimentação espúria contida na imagem.

Figura 4.4: Deslocamento dos frames na direção x antes e depois da estabilização.

Figura 4.5: Deslocamento dos frames na direção y antes e depois da estabilização.

Figura 4.6: Quimografia realizada antes (a) e após da estabilização do vídeo (b).

4.3 *Optical Flow*

Com a técnica de *Optical Flow* podemos observar que não é muito eficaz na estimação dos movimentos dos vídeos. As figuras 4.7 e 4.8 mostram os gráficos comparando os movimentos dos *frames* antes e após a estabilização nos eixos X e Y respectivamente e a tabela 4.3 mostra a variância dos movimentos antes e depois da estabilização.

Tabela 4.3: Variância dos deslocamentos antes a depois da estabilização utilizando a técnica de *Optical Flow*.

	X antes	X depois	Y antes	Y depois	Varição X (%)	Varição Y (%)
Variância	26.92	28.40	34.88	5.49	5.49	-84.26

Após o processo de estabilização foi feita a quimografia a partir do vídeo. As figuras 4.9 (a) e 4.9 (b) mostram as quimografias realizadas antes a após a estabilização do vídeo.

Figura 4.7: Movimento antes e após a estabilização no eixo X utilizando *Optical Flow*.

Figura 4.8: Movimento antes e após a estabilização no eixo Y utilizando *Optical Flow*.

Figura 4.9: Quimografia realizada antes (a) e após da estabilização do vídeo (b).

4.4 Discussão

A tabela 4.4 mostra como a variância dos movimentos nos eixos x e y antes e após a estabilização variou, utilizando cada uma das técnicas avaliadas. Nota-se que cada um dos métodos apresentaram resultados bastante distintos.

Tabela 4.4: Comparação da variância entre os três métodos CE - Correlação espacial, CF - Correlação de Fase e OF - *Optical Flow*.

	X antes	X depois	Y antes	Y depois	Variação X (%)	Variação Y (%)
CE	16.44	0.36	65.61	0.05	-97.83	-99.92
CF	6.82	11,56	11.18	8.40	69.50	-24,86
OF	26.92	28.40	34.88	5.49	5.49	-84.26

De um modo geral o método de correlação de espacial foi o que apresentou melhor resultado, pode-se observar na figura 4.1 o movimento estimado antes da estabilização no eixo x variou entre 0 a 16 pixels e após o processamento ficou entre -2 a 1 pixels, uma redução de 97%. No eixo y , figura 4.2, o movimento estimado antes da estabilização variou de 0 a 23 pixels e após o processamento ficou entre 0 e 1 pixel, uma redução de 99%, mostrando que esse método foi bastante eficaz para eliminar os movimentos do vídeo.

O método de correlação de fase não mostrou melhoras significativas na estabilização, analisando a figura 4.4 observamos que antes da estabilização no eixo x o movimento

estimado variou de -6 a 6 pixels e após o processamento o mesmo foi de -4 a 6 pixels, mas o aumento de 69% na variância não significa muita degradação da imagem nessa direção. Na figura 4.5 o movimento estimado antes da estabilização no eixo y foi de -1 a 9 pixels e após o processamento variou de -1 a 7, uma redução de 24% que não representou melhora significativa na qualidade da imagem.

Com a técnica de *Optical Flow* pode-se observar que a estimação do movimento no eixo x , figura 4.7, variou de -5 a 15 e -5 a 10 pixels antes e após a estabilização respectivamente, resultando em um aumento de 5% na variância do movimento. No eixo y , figura 4.8, a variação foi de -15 a 1 e -1 a 8 pixels antes e após a estabilização respectivamente. Uma redução de 84%, que da mesma forma que a correlação de fase não foi significativa para a melhoras da imagem; já que essas duas últimas técnicas não conseguiram estimar corretamente os movimentos contidos no vídeo.

Capítulo 5

Conclusões

Com os resultados encontrados, pode se observar que os métodos utilizados para a estabilização dos vídeos apresentaram desempenhos diferentes entre si.

O método de correlação espacial foi o que apresentou o melhor resultado entre eles, capaz de fazer uma correta estimação dos movimentos e corrigí-los. Entretanto, esse método apresenta uma dificuldade em ser utilizado, já que diferentes escolhas da região e tamanho do *template* podem apresentar resultados diferentes.

A correlação de fase apresentou, sob certas condições de movimentos não abruptos da câmera, resultados satisfatórios para a compensação do movimento. Um fator complicador do método consiste no fato de que as imagens da laringe humana, em muitos casos, possuem pouco contraste, principalmente quando as pregas vocais se encontram fechadas. A falta de sincronismo entre a emissão da luz e a captura da imagem também é uma fonte de erro na estimação do movimento, pois introduz *frames* totalmente escuros.

A técnica de *Optical Flow* não demonstrou desempenho satisfatório, chegando até mesmo degradar a imagem original. A provável causa desse problema pode ser o pouco contraste presente nas sub-regiões escolhidas bem como o problema com a iluminação não ser constante ao longo do tempo, o que ocasiona erros de estimação do movimento.

Fazendo a equalização de histograma, em certos casos, conseguimos uma melhora significativa no processo de estabilização, principalmente quando utilizado com a correlação espacial.

5.1 Trabalhos futuros

Em trabalhos futuros pode-se utilizar outros algoritmos para fazer a estabilização dos vídeos, como o *Three-Step-Search*, algoritmo *Piramidial* (Balakirsky e Chellappa, 1996), outras técnicas de *Optical Flow* como a de Lukas-Kanade e *Block-Match*. Implementações no método de correlação espacial podem ser realizadas, como um modo para localizar a melhor posição de extração do *template* e escolha do seu tamanho em função do tamanho da imagem.

Uma padronização no sistema de captura deve ser feita, de modo que todos os vídeos possuam um mesmo tamanho, 320x240 pixel por exemplo, e melhoras no sincronismo entre a emissão da luz do estrobo e a captura da imagem, dessa maneira eliminando os *frames* escuros presentes no *frame*.

A utilização de filtros, de Kalman, LMS (Erturk, 2002; Osbek e Erturk, 2004) ou *wavelets* por exemplo, para a melhora da estimação dos movimentos e atenuação e/ou eliminação dos ruídos nas imagens pode ser avaliados.

Métodos para correção de movimentos de rotação ou mudança de escala, bem como um sistema de estabilização de imagens em tempo real poderão ser implementados.

REFERÊNCIAS

- ANATOMY & Physiology of the Larynx. (1999). Disponível em: <<http://www.gbmc.org/voice/anatomyphysiologyofthelarynx.cfm>>. Acesso em: 1 June 2007.
- BALAKIRSKY, S.B.; CHELLAPPA, R. (1996). Performance characterization of image stabilization algorithms. **Real-Time Imaging**, London, v.2, n.5, p.297–313, Oct.
- CENSI, A.; FUSIELLO, A.; ROBERTO, V. (1999). Image stabilization by features tracking, In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, 10., 1999, Veneza. **Proceedings...** New York: IEEE. p.665-672.
- DE CASTRO, E.; MORANDI, C. (1987). Resgistration of traslated and rotated images using finite fourier transform. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v.9, n.5, p.700–703.
- ERTURK, S. (2002). Real time digital image stabilization using kalman filters. **Real Time Imaging**, London, v.8, n.4, p.317–328, Aug.
- _____. (2003). Digital image stabilization with sub-image phase correlation based global motion estimation. **IEEE Transactions on Consumer Eletronics**, New York, v.49, n.4, p.1320–1325, Nov.
- ERTURK, S.; DENNIS, T.J. (2000). Image sequence stabilization based on DFT filtering. **IEEE Proceidings. Vision Image Signal Processing**, New York, v.147, n.2, p.95–102.
- EYSHOLDT, U.; ROSANOWSKI, F.; HOPPE, U. (2003). Vocal fold vibration irregularities caused by different types of laryngeal asymmetry. **European Archives Otorhinolaryngology**, Berlin, v.260, n.8, p.412–417, Sept.
- FEDOROV, D. (2002). **Sistema semi-automático de registro e mosaico de imagens**. Tese (Mestrado) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2002.
- GONZALEZ, R.C.; WOODS, R.E. (1993). **Digital image processing**. New York: Addison Wesley.

HERTEGARD, S.; LARSSON, H.; WITTENBERG, T. (2003). High speed imaging: applications and development. **Logopedics Phoniatrics Vocology**, Oslo, v.28, n.3, p.133–139.

HORN, B.K.P.; SCHUNK, B.G. (1981). Determining optical flow. **Artificial Intelligence**, Amsterdam, v.17, n.1/3, p.185-203, Aug.

LAMBERT INSTRUMENTS (2007). Disponível em: <<http://www.lambert-instruments.com/upload/camera-for-Videokymography-KymoCAM.pdf>>. Acesso em: 1 July 2007.

ISHIZAKA, K.; MATSUDARIA, M. (1972). **Fluid mechanical considerations of vocal cord vibration monograph 8, speech communication res lab 1972**. Santa Barbara. Technical Report.

JAHNE, B. (1997). **Digital image processing: Concepts, algorithms and scientific applications**. 4thed. Berlin: Springer Verlag.

JAIN, R.; CASTURY, R.; SHUNK, B.G. (1995). **Machine vision**. New York: McGraw-Hill.

KIM, D. et al. (2003). Videostrobokymographic analysis of benign vocal fold lesion. **Acta Otol-Lryngologica**, Stokholm, v.123, n.9, p.1102–1109.

KO, S.J.; LEE, S.H.; LEE, K.H. (1998). Digital image stabilizing algorithms based on bit-plane matching. **IEEE Transactions on Consumer Electronics**, New York, v.44, n.3, p.617–622, Aug.

KOISHI, H.U. et al. (2003). Variação da intensidade vocal: estudo da vibração das pregas vocais em seres humanos com videoquimografia. **Revista Brasileira de Otorrinolaringologia**, São Paulo, v.69, n.4, p.464–470, jul./ago.

KUGLIN, C.D.; HINES, D.C. (1975). The Phase correlation image alignment method. In: INTERNATIONAL CONFERENCE ON CYBERNETICS AND SOCIETY, 1975, Bucharest. **Proceedings...** New York: IEEE. P.163-165.

LEE., J.S. et al. (2001). A method for assessing the regional vibratory pattern of vocal folds by analysing the video recording of stroboscopy. **Medical and Biological Engineering and Computing**, New York, v.39, n.3, p.273–278, May.

LEHMANN, T.M. et al. (1996). A Rotation-extended cepstrum technique and its application to medical images. **Pattern Recognition and Image Analysis**, Moscow, v.6, n.3, p.592–604.

LIPTON, A.J.; FUJIYOSHI, H.; PATIL, R.S. (1998). Moving target classification and tracking from real-time video. In: IEEE WORKSHOP ON APPLICATIONS OF COMPUTER VISION, 4., 1998, Princeton. **Proceedings...** New York: IEEE. p.8–14.

LUCAS, B.D.; KANADE, T. (1981). An Iterative image registration technique with an application to stereo vision. In: DARPA IMAGE UNDERSTANDING WORKSHOP, 1981, Washington. **Proceedings...** New York: IEEE. p.121–130.

MONTAGNOLI, A.N. (2006). **Cirurgia virtual da laringe**. 125p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

OHYAMA, N. et al. (1987). Compensation of motion blur in CCD color endoscope images. **Applied Optics**, Easton, v.26, n.5, p.909–912, Mar.

OPEN SOURCE COMPUTER VISION LIBRARY. (2007). Disponível em:<<http://sourceforge.net/projects/opencvlibrary/>>. Acesso em: 10 Mar. 2007.

OSBEK, F.; ERTURK, S. (2004). Image sequence stabilization based on adaptive polynomial filtering using LMS and RLS algorithms. **Journal of Electrical & Electronics Engineering**, Istanbul, v.4, n.2, p.1233–1242.

OSHIMA, M. et al. (1989). VHS camcorder with electronic image stabilizer. **IEEE Transactions on Consumer Electronics**, New York, v.35, n.4, p.749–757.

PEARSON JR., J.J. et al. (1977). Video-rate image correlation processor. In: TESCHER, A.G. (Ed.). **Application of digital image processing**. Bellingham: SPIE. p.197–205. (SPIE, 119).

REDDY, B.S.; CHATTERJI, B.N. (1996). An FFT based technique for translation, rotation and scale-invariant image registration. **IEEE Transactions on Image Processing**, New York, v.5, n.8, p.1266–1271, Aug.

SANGWINE, S.J.; MOXEY, C.E. (2001). Vector phase correlation. **Electronics Letters**, London, v.37, n.25, p.1513–1515, Dec.

SATO, K. et al. (1993). Control techniques for optical image stabilizing system. **IEEE Transactions on Consumer Electronics**, New York, v.39, n.3, p.461–466, June.

SCHALKOFF, R.J. (1989). **Digital image processing and computer vision**. New York: John Wiley.

SHI, J.; TOMASI, C. (1994). Good features to track. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 1994, Seattle. **Proceedings...** New York: IEEE. p.593-600.

STEVENS, K. (1977). Physics of laryngeal behavior and larynx modes. **Folia Phonetica**, Basel, v.34, n.4, p.266–277.

SUNG, M.W. et al. (1999). Videostrobokymography: a new method for the quantitative analysis of vocal fold vibration. **The Laryngoscope**, Saint Louis, v.109, n.11, p.1859–1863, Nov.

TEKALP, A.M. (1995). **Digital video processing**. Upper Saddle River: Prentice Hall.

VASEGHI, S.V. (2001). **Advanced digital signal processing and noise reduction**. 2nded. New York: John Wiley.

VELLA, F. et al. (2002). Digital image stabilization by adaptive motion vectors filtering. **IEEE Transactions on Consumer Electronics**, New York, v.48, n.3, p.796–801, Aug.

SVEC, J.G. (2000). **On Vibration properties of human vocal folds**: voice registers, bifurcations, resonance characteristics, development and application of videokymography. PhD. Thesis - University of Groningen, 2000. Disponível em: <<http://irs.ub.rug.nl/ppn/240208714>>. Acesso em: 10 June 2007.

SVEC, J.G.; SRAM, F.; SCHUTTE, H.K. (2000). Development and application of videokymography for high-speed examination of vocal-fold vibration. In: LINGUISTICS AND PHONETICS CONFERENCE, 2000, Prague. **Proceedings...** Prague: Karolinum. p.3–10.

SVEC, J.G.; SCHUTTE, H.K. (1996). Videokymography: high-speed line scanning of vocal fold vibration. **Journal of Voice**, Saint Louis, v.10, n.2, p.201–205.

YENI, A.A.; ERTURK, S. (2005). Fast digital image stabilization using one bit transform based sub-image motion estimation. **IEEE Transactions on Consumer Electronics**, New York, v.51, n.3, p.917–921, Aug.