

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**CARLOS ALBERTO RAMIREZ BEHAINE**

**RECONHECIMENTO DE PADRÕES  
COM TRATAMENTO DE INCERTEZAS  
NA LOCALIZAÇÃO DE MARCADORES  
E MODELOS ATIVOS DE FORMAS**

Porto Alegre  
2013

**CARLOS ALBERTO RAMIREZ BEHAINE**

**RECONHECIMENTO DE PADRÕES  
COM TRATAMENTO DE INCERTEZAS  
NA LOCALIZAÇÃO DE MARCADORES  
E MODELOS ATIVOS DE FORMAS**

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Engenharia de Computação

ORIENTADOR: Prof. Dr. Jacob Scharcanski

Porto Alegre  
2013

**CARLOS ALBERTO RAMIREZ BEHAINE**

**RECONHECIMENTO DE PADRÕES  
COM TRATAMENTO DE INCERTEZAS  
NA LOCALIZAÇÃO DE MARCADORES  
E MODELOS ATIVOS DE FORMAS**

Esta tese foi julgada adequada para a obtenção do título de Doutor em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: \_\_\_\_\_  
Prof. Dr. Jacob Scharcanski , UFRGS  
Doutor pela Universidade de Waterloo – Waterloo, Canadá.

Banca Examinadora:

Prof. Dra. Maria Cristina Ferreira de Oliveira, ICMC-USP  
Doutor pela University of Wales – Bangor, UK

Prof. Dr. Luciano Silva da Silva, FURG  
Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Prof. Dr. Wilson Pires Gavião Neto, UNIRITTER  
Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Prof. Dr. Valner João Brusamarello, PPGEE-UFRGS  
Doutor pela Universidade Federal de Santa Catarina – Florianópolis, Brasil

Prof. Dr. Walter Fetter Lages, PPGEE-UFRGS  
Doutor pelo Instituto Tecnológico de Aeronáutica – São José dos Campos, Brasil

Coordenador do PPGEE: \_\_\_\_\_  
Prof. Dr. João Manoel Gomes da Silva Junior, PPGEE-UFRGS

Porto Alegre, outubro de 2013.

## **DEDICATÓRIA**

À menina linda, pelo mês maravilhoso.

## AGRADECIMENTOS

Ao professor Jacob Scharcanski pela orientação durante o curso de doutorado.

Ao professor Arturo Suman Bretas pelo suporte durante o estágio de docência em proteção e estabilidade dos sistemas elétricos.

Ao professor Valner João Brusamarello pelo suporte durante o estágio de docência no laboratório de circuitos elétricos I-B.

Ao Flavio Makishi por me emprestar a câmera digital profissional.

Ao pessoal do laboratório de circuitos analógicos pela ajuda com os equipamentos.

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela ajuda econômica para a publicação da minha produção científica e pelo formato deste documento em  $\text{\LaTeX}$ .

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES, pela provisão da bolsa de doutorado.

## RESUMO

As imagens são sinais que possuem muita informação. Os objetos representados em imagens podem sofrer deformações fazendo com que suas características mudem, o que dificulta o reconhecimento do objeto. A chave para o sucesso de um sistema de reconhecimento de padrões em imagens é escolher adequadamente a sua abordagem e um modelo para as feições presentes nas imagens. Uma das dificuldades é extrair e selecionar as feições que são mais discriminantes entre as diferentes classes usadas para modelar um objeto. Os modelos ativos de formas (*Active Shape Models* ASM) adaptam-se às deformações de um objeto. O objeto a ser modelado, pode ser representado com um modelo de pontos distribuídos (*Point Distribution Models* PDM). O PDM consiste de pontos de interesse ou marcadores, que permitem a extração de feições em localizações específicas do objeto. Após tratar a incerteza da localização e oclusão dos marcadores é possível extrair as feições mais representativas, obtendo-se um desempenho alto em termos da taxa de reconhecimento. Nesta tese são introduzidas novas formas para extrair e selecionar feições com modelos ativos de formas, que melhoram a taxa de classificação onde há objetos deformáveis. Esta tese é inovadora no sentido de aperfeiçoar o uso de ASMs na classificação de faces humanas, e na sua aplicação no monitoramento visual de outros tipos de objetos deformáveis.

**Palavras-chave:** Reconhecimento de Padrões em Imagens, Visão Computacional, Modelos Ativos de Formas (ASM), Monitoramento de Objetos Deformáveis.

## **ABSTRACT**

Images are signals that have a lot of information. The objects depicted in images may suffer deformations causing changes in their characteristics, which hinders the recognition of the object. The key to the success of a system of pattern recognition in images is to choose properly your approach and a model for the features in the images. One difficulty is to extract and select the features that are most discriminating between different classes used to model an object. The Active Shape Models (ASM) adapt to deformations of an object. The object to be modeled can be represented with a Points Distribution Model (PDM). The PDM consists of points of interest or landmarks that allow the extraction of features in specific locations of the object. After treating the uncertainty of the location and occlusion of the landmarks it is possible to extract the most representative features, obtaining a high performance in terms of recognition rate. This thesis introduces new ways to extract and select features with ASMs, which improve the classification rate where deformable objects are present. This thesis is innovative in the sense that improves the use of ASMs in the classification of human faces, and that can be applied in visual monitoring of other types of deformable objects.

**Keywords: Image Pattern Recognition, Computer Vision, Active Shape Models (ASM), Monitoring of Deformable Objects.**

# SUMÁRIO

<b>LISTA DE ILUSTRAÇÕES</b>	9
<b>LISTA DE TABELAS</b>	11
<b>LISTA DE ABREVIATURAS</b>	12
<b>LISTA DE SÍMBOLOS</b>	14
<b>1 INTRODUÇÃO</b>	16
1.1 Prólogo	16
1.2 Motivação	19
1.3 Objetivos	19
1.4 Contribuições	20
1.5 Estrutura do Documento	20
<b>2 MODELOS ATIVOS DE FORMAS</b>	22
2.1 Revisão Bibliográfica	22
2.2 Teoria dos ASMs	24
2.3 Seleção de Feições	27
<b>3 INOVAÇÕES PROPOSTAS NOS MODELOS ATIVOS DE FORMAS</b>	31
3.1 Modelos Ativos de Formas e Reconhecimento de Objetos	32
3.1.1 Composição dos Vetores de Feições com Marcadores Chave em ASMs	34
3.1.2 Seleção de Feições Consistentes para Reconhecimento de Objetos com ASMs	35
3.2 Modelos Ativos de Formas Modificados	39
3.2.1 Marcadores para a Modelagem das Deformações de um Objeto	39
3.2.2 Obtenção de Modelos de Aparência Utilizando Perfis RGB	41
3.2.3 Localização dos Marcadores Utilizando os Perfis RGB	44
<b>4 RESULTADOS EXPERIMENTAIS E DISCUSSÃO</b>	47
4.1 Aplicações em Reconhecimento de Faces	47
4.2 Aplicações em Monitoramento Visual de Painéis	50
4.3 Aplicações em Monitoramento Visual de Multipainéis	59
4.4 Discussão	67
<b>5 CONCLUSÕES</b>	69
<b>REFERÊNCIAS</b>	72



<b>APÊNDICE A</b>	<b>INFORMAÇÃO MÚTUA AJUSTADA . . . . .</b>	<b>78</b>
<b>APÊNDICE B</b>	<b>ESTIMATIVA DO <math>\alpha</math> ÓTIMO . . . . .</b>	<b>80</b>

## LISTA DE ILUSTRAÇÕES

Figura 1:	Modos e etapas do esquema clássico do reconhecimento de padrões com abordagem estatística. . . . .	17
Figura 2:	(a) Ilustração de uma imagem de uma face ( $250 \times 250 \text{ pixels}$ ) da base de dados Essex (ESSEX, 2011), e a sua (b) Representação em um PDM de $N = 68$ marcadores. . . . .	18
Figura 3:	Ilustração da localização das normais nos $N = 12$ marcadores de um PDM de um objeto. . . . .	26
Figura 4:	Ilustração de: (a) marcadores $p_i$ (PDM) utilizado para modelar uma face no ASM; e (b) Localização de marcadores em uma imagem da face. . . . .	34
Figura 5:	Representação de um PDM de 12 marcadores com a (a) Ilustração dos pesos discriminantes $a_i$ nos marcadores, e a (b) Ilustração da seleção de $R = 5$ marcadores (em vermelho) para a classificação de um objeto. . . . .	37
Figura 6:	Ilustração de um PDM utilizando marcadores identificados $P_{a_n}$ (círculos vermelhos vazios), e os marcadores interpolados (círculos azuis) que modelam um ponteiro de um painel analógico utilizando um ASM com $N = 9$ , $L = 8$ e $Q = 17$ . . . . .	41
Figura 7:	Ilustração de um espaço para a localização do $q$ -ésimo marcador onde uma busca local segue a direção normal. . . . .	46
Figura 8:	Ilustração de $R = 27$ marcadores $p_i$ (PDM) selecionados em vermelho no teste do método. . . . .	49
Figura 9:	Comparação de desempenho utilizando: (a) Avaliação do método; e (b) Avaliação sistêmica. . . . .	50
Figura 10:	(a) Esquema de redundância ativa para monitoramento visual de medidores com painel analógico, e (b) Rede utilizada para o monitoramento de medidores. . . . .	51
Figura 11:	(a) Ilustração de marcadores $p_i$ (PDM) utilizados no ASM proposto para o painel analógico, com $N = 14$ e $R = 3$ marcadores principais localizados para a ROI, e (b) Outra forma do PDM em um ASM para um painel analógico. . . . .	52
Figura 12:	(a) Exemplo de valores da função de pesos $a_i$ com $R_e = 5$ no PDM de um painel analógico (marcadores selecionados em vermelho para um valor $R_e = 5$ ), e a sua (b) Ilustração no PDM respectivo (marcadores selecionados em círculos vermelhos vazios para um valor $R_e = 5$ ). . . . .	53
Figura 13:	(a) Painel analógico do medidor HIOKI MODEL 3007, e (b) Comparação do ASM para o painel analógico do medidor HIOKI MODEL 3007, marcadores em pontos vermelhos e bordas em pontos azuis. . . . .	55

Figura 14:	Média do erro da comparação ASM por <i>pixel</i> obtido no painel analógico, (a) Modelo classe $k = 1$ , (b) Modelo classe $k = 2$ , e (c) Modelo classe $k = 3$ . . . . .	56
Figura 15:	Medições do método do monitoramento automático visual apresentado (Volts). . . . .	57
Figura 16:	Comparação do ASM em um painel analógico com um PDM ilustrado na Figura 11 (b), marcadores em pontos vermelhos e bordas em pontos azuis. . . . .	59
Figura 17:	(a) Ilustração de um PDM utilizando marcadores identificados, (círculos vazios em vermelho) e marcadores interpolados (círculos em azul) para modelar um ponteiro analógico ASM, neste caso, $N = 6$ , $L = 5$ e $Q = 11$ , (b) Outro caso de PDM com $N = 11$ , $L = 10$ e $Q = 21$ . . . . .	61
Figura 18:	Ilustração de um diagrama de blocos de uma rede para monitoramento visual automático utilizando concorrentemente ASM para $U$ painéis analógicos. . . . .	62
Figura 19:	(a) Ilustração de uma imagem de multipainéis analógicos: (1) modelo RENZ BRASIL (AC V), (2) modelo MF133 (AC V) e (3) modelo RENZ BRASIL (AC mA); e (b) Resultados da comparação ASM com $\epsilon = 6$ <i>pixels</i> , depois de 10 iterações respectivamente. . . . .	64

## LISTA DE TABELAS

Tabela 1:	Comparação de desempenho para o conjunto de testes (maiores valores de exatidão) . . . . .	48
Tabela 2:	Estatística da exatidão para o esquema apresentado de monitoramento visual de medidores com painel, (a) 3 <i>pixels</i> por $\checkmark$ e (b) 9 <i>pixels</i> por $\checkmark$ . . . . .	58
Tabela 3:	Estatística da exatidão para o esquema apresentado de monitoramento visual do painel ilustrado na Figura 16, 5 <i>pixels</i> por $\checkmark$ . . . . .	58
Tabela 4:	Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (1). . . . .	66
Tabela 5:	Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (2). . . . .	66
Tabela 6:	Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (3). . . . .	66
Tabela 7:	Desempenho comparativo do monitoramento visual automático proposto para multipainéis analógicos de medidores, ilustrado para um sistema com $U = 3$ painéis. . . . .	67

## LISTA DE ABREVIATURAS

2D	Duas dimensões
3D	Três dimensões
AAM	Modelos de Aparências Ativas ( <i>Active Appearance Models</i> )
AC	Corrente Alternada ( <i>Alternating Current</i> )
AMI	Informação Mútua Ajustada ( <i>Adjusted Mutual Information</i> )
ASM	Modelos Ativos de Formas ( <i>Active Shape Models</i> )
BMP	<i>Windows Bitmap</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DC	Corrente Direta ( <i>Direct Current</i> )
DFBFR	Fusão Melhorada de Dados para Reconhecimento de Faces ( <i>Data Fusion Boosted Face Recognition</i> )
GA	Algoritmos Genéticos ( <i>Genetic Algorithms</i> )
GMM	Modelo de Mistura de Gaussianas ( <i>Gaussian Mixture Model</i> )
IBM-PC	<i>Personal Computer</i>
JPEG	<i>Joint Photographic Experts Group</i>
Matlab	<i>Matrix Laboratory</i>
MI	Informação Mútua ( <i>Mutual Information</i> )
PCA	Análise de Componentes Principais ( <i>Principal Component Analysis</i> )
PDM	Modelo de Pontos Distribuídos ( <i>Point Distribution Model</i> )
PPGEE	Programa de Pós-Graduação em Engenharia Elétrica
RAM	<i>Random Access Memory</i>
RGB	Modelo de cor <i>Red Green Blue</i>
ROI	Região de Interesse ( <i>Region of Interest</i> )
SBS	<i>Sequential Backward Selection</i>
SBFS	<i>Sequential Backward Floating Search</i>
SFFS	<i>Sequential Forward Floating Search</i>

- SFS *Sequential Forward Selection*
- SRDA *Análise de Regressão Discriminante Espectral (Spectral Regression Discriminant Analysis)*
- SVM *Máquinas de Vetores de Suporte (Support Vector Machines)*

## LISTA DE SÍMBOLOS

$\Sigma$	Somatório
$J$	Número de imagens de treinamento por classe
$M$	Número de feições
$K$	Número de classes
$p_i$	$i$ -ésimo marcador
$S_k$	Representação do PDM do objeto da $k$ -ésima classe
$\mu_{F_{m,i}}$	Média das medições na imagem da $m$ -ésima feição do $i$ -ésimo marcador em todas as $J$ amostras das imagens de treinamento
$\sigma_{F_{m,i}}^2$	Variância das medições na imagem da $m$ -ésima feição do $i$ -ésimo marcador em todas as $J$ amostras das imagens de treinamento
$\mathbf{u}_{F_{m,i}}$	Vetor da $m$ -ésima feição do $i$ -ésimo marcador das classes
$a_i$	Função de peso para o $i$ -ésimo marcador
$G_{m,i}^k$	Gaussiana que representa a classe de face $k$ , de média $\mu_{F_{m,i}}^k$ e variância $\sigma_{F_{m,i}}^2$
$\top$	Operador de transposição de matrizes ou vetores
$A(\mathbf{u}, \mathbf{v})$	AMI entre os vetores $\mathbf{u}$ e $\mathbf{v}$
$\hat{H}$	Entropia
$MI$	Informação mútua
$B_i^k$	Confiabilidade das feições medidas no $p_i$ na classe $k$ -ésima
$\alpha$	Valor escalar entre 0 e 1.
$\Omega_{ck}$	Modelo de Gaussianas e pesos para a classe de face $k$
$U$	Número de painéis
$\{P_{ja_n}\}$	marcadores identificados da $j$ -ésima imagem de treinamento, $n = 1, \dots, N$
$\{P_{jl_b}\}$	marcadores interpolados da $j$ -ésima imagem de treinamento, $l = 1, \dots, L$
$\mathbf{X}$	Matriz de treinamento dos PDMs
$\mathbf{x}$	Modelo das deformações
$\mathcal{L}$	Luminância

$\mathbf{g}_{ji_{\mathcal{L}}}$	Vetor derivada do perfil normalizado da $j$ -ésima imagem de treinamento no $i$ -ésimo marcador
$\mathbf{G}_{i_{\mathcal{L}}}$	Matriz de perfis de luminância no $i$ -ésimo marcador
$\mathbf{g}_{i_{\mathcal{L}}}^t$	Vetor derivada do perfil normalizado da $t$ -ésima imagem de teste no $i$ -ésimo marcador
$M_d(\cdot)$	Distância de Mahalanobis
$\hat{\mathbf{g}}_q$	Modelo do perfil no $q$ -ésimo marcador
$\mathbf{G}_q$	Matriz de perfis no $q$ -ésimo marcador.
$\hat{\mathbf{g}}$	Modelo de aparência ou dos perfis
$\ \cdot\ _2$	Norma Euclidiana
$\mathbf{g}_q^{t(\varrho)}$	Vetor derivada do perfil normalizado de uma $t$ -ésima imagem de teste no $q$ -ésimo marcador, deslocado $\varrho$ posições ao longo do intervalo $[-\eta, \eta]$
$\Delta_{\mathbf{x}_z}$	Vetor de deslocamento em uma $z$ -ésima instância do modelo das deformações
$T_{\tau, \xi, \theta}(\cdot)$	Operador de rotação $\theta$ , de escala $\xi$ e translação de coordenadas $(x, y)_{\tau}$
$\check{\epsilon}$	Escala mínima de divisão de um painel analógico
$\delta$	Tolerância de escala de um painel analógico
$\mu$	Média do erro absoluto
$\sigma$	Desvio padrão do erro absoluto
$\Omega$	Média do posto dos métodos ASMs para $U$ painéis analógicos
$PPV$	Valor Preditivo Positivo
$TP$	Verdadeiro Positivo
$FP$	Falso Positivo
$\mathbb{R}$	Conjunto dos números reais.



# 1 INTRODUÇÃO

## 1.1 Prólogo

O objetivo principal de um sistema de reconhecimento de padrões é fornecer uma classificação. Em um esquema clássico de reconhecimento de padrões, a classificação é suportada por uma etapa de aprendizado (JAIN; DUIN; MAO, 2000). Um sistema de reconhecimento de padrões tem dois modos de operação: o modo de treinamento e o modo de teste. O modo de treinamento realiza processos de seleção e aprendizagem de sinais para fornecer suporte e estimativas adequadas para a etapa de classificação. O modo de treinamento possui três etapas: um pré-processamento, uma extração e seleção de características e um aprendizado. O modo de teste é formado por três etapas essenciais: um pré-processamento, uma medição de características e uma classificação. Um esquema clássico de reconhecimento de padrões com abordagem estatística é ilustrado na Figura 1 (JAIN; DUIN; MAO, 2000). As linhas na Figura 1 representam relações sequenciais entre etapas que pertencem ao mesmo modo e as linhas pontilhadas representam relações funcionais entre etapas que não pertencem ao mesmo modo.

O surgimento dos modelos gráficos estatísticos, entre eles os modelos ativos de formas (ASM) (COOTES et al., 1995), retomaram novamente o interesse da pesquisa com abordagem clássica estatística. Mais recentemente, o foco da atenção gira em torno de dois fatos importantes: a maior dificuldade para a descrição das imagens é extrair as feições que são mais consistentes, e que a chave para o sucesso de um sistema de reconhecimento de padrões em imagens é escolher adequadamente a sua abordagem e o modelo adequado para as feições presentes nas imagens (ZHENG et al., 2008).

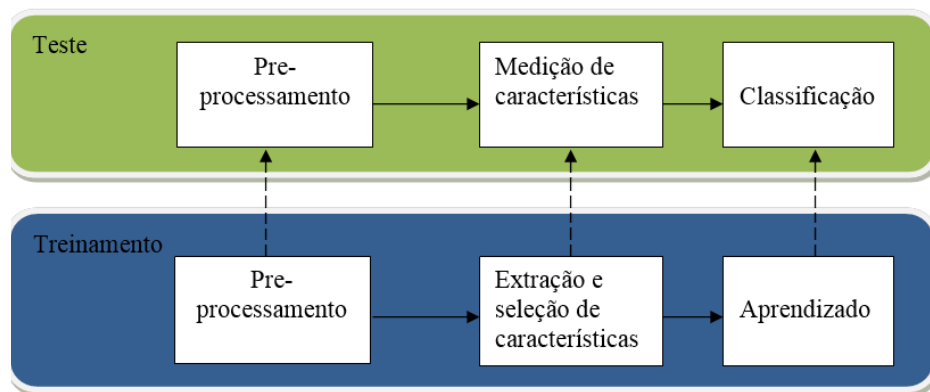


Figura 1: Modos e etapas do esquema clássico do reconhecimento de padrões com abordagem estatística.

O reconhecimento de padrões em imagens envolve sinais cuja característica principal é o alto conteúdo de informação. O reconhecimento de padrões deformáveis em imagens com aplicação em objetos pretende obter uma extração e classificação adequada do objeto modelado. Os modelos ativos de formas (ASM) representam deformações espaciais de um objeto. Os pontos de interesse do objeto a ser modelado no ASM são chamados marcadores. Os marcadores são representações discretas espaciais do objeto. Uma quantidade de  $N$  marcadores formam um conjunto de pontos chamado modelo de pontos distribuídos ou (PDM). Na Figura 2 (a) e na Figura 2 (b) ilustra-se uma imagem de uma face e o seu PDM de  $N = 68$  marcadores, respectivamente. O número  $N$  de marcadores é minúsculo quando comparado ao número de *pixels* utilizados em uma representação densa do objeto presente na imagem, por exemplo a Figura 2 (a) utiliza  $250 \times 250 = 62500$  *pixels*, entretanto a Figura 2 (b) utiliza somente 68 pontos em anexo ou marcadores. Pode-se dizer então, que o PDM fornece uma representação reduzida em termos de marcadores utilizados, quando comparado com representações densas da imagem de um objeto.

Os ASMs encaixam dentro de uma parte das técnicas de visão computacional que pretendem ter uma solução visual a problemas de reconhecimento de objetos que sofrem deformações e oclusões (COOTES et al., 1995). Os painéis analógicos são objetos que sofrem deformações para representar grandezas. Os painéis analógicos podem ser monitorados visualmente, mas estes estão normalmente em presença de mudanças de lu-



Figura 2: (a) Ilustração de uma imagem de uma face (  $250 \times 250 \text{ pixels}$  ) da base de dados Essex (ESSEX, 2011), e a sua (b) Representação em um PDM de  $N = 68$  marcadores.

minosidade, sombras, não uniformidade do fundo, escalas diferentes, etc. A habilidade humana para monitorar tais painéis simultaneamente, diminui com o número de painéis (SCHAFFER; GOULD, 1966). O monitoramento visual automático de painéis analógicos remotos é uma técnica de medição sem contato baseado em visão computacional, que é adequada para a leitura e monitoramento de instrumentos remotos que não tem acesso fácil através de um enlace de comunicação, ou na implementação de laços redundantes de monitoramento de instrumentos. Abordagens do monitoramento visual remoto têm sido utilizadas em vários problemas práticos, tais como leitura de medidores de energia (ZHAO et al., 2005), calibração automática remota de painéis (ALEGRIA; SERRA, 2000), teste automático remoto (GANESH; THANUSHKODI, 2006) e controle de qualidade (LIU; WU; ZOU, 2010). Os fundamentos de correlação e comparação de *templates* têm sido muito utilizados em aplicações de monitoramento visual remoto, entretanto a utilização destes bem conhecidos métodos de processamento de sinais e imagens, frequentemente consomem muito tempo e são susceptíveis a erros ( i.e. peculiaridades da imagem, ou deficiência de um contraste local adequado) (ZHAO et al., 2005; GANESH; THANUSHKODI, 2006).

O escopo geral desta tese é o reconhecimento de padrões em imagens utilizando modelos ativos de formas, especificamente na extração e seleção de características, na medição de características e na classificação, cuja motivação é mostrada a seguir.

## 1.2 Motivação

As aplicações do ASM em imagens são numerosas, por exemplo a segmentação do cérebro em imagens de ressonância magnética (DUTA; SONKA, 1998), a segmentação de faces em tempo real (FROBA et al., 2001), o alinhamento de faces (WANG et al., 2002), a estimativa de pose da face (YARI; CAVALCANTI; SCHARCANSKI, 2010; YARI; SCHARCANSKI, 2011), etc. As feições biométricas em sistemas de reconhecimento de faces são uma das mais confiáveis e menos invasivas alternativas para a autenticação da identidade das pessoas (JENKINS; BURTON, 2008). A classificação de faces é ainda uma tarefa desafiadora. As faces representadas em imagens podem sofrer deformações fazendo com que suas características mudem, o que dificulta o reconhecimento da face. O ASM é uma técnica de correspondência ou comparação adaptativa de formas, que tem sido utilizada frequentemente para localizar feições faciais em imagens de faces através de marcadores. Os erros da comparação da face com ASM podem ser altos em algumas localidades da face (HILL; COOTES; TAYLOR, 1996), apesar de melhoras recentes feitas na técnica do ASM (WAN; LAM; NG, 2005; JUNMO; MüJDAT; ALAN, 2007). A detecção inconsistente de feições faciais acontece frequentemente porque o ASM treinado converge em direção às bordas salientes da imagem, e se estas bordas salientes forem ruidosas ou distorcidas pelo sombreamento ou iluminação, podem acontecer comparações erradas de feições. Estes tipos de erros têm limitado a utilização do ASM para tarefas em reconhecimento de faces e tarefas em reconhecimento de objetos em geral.

## 1.3 Objetivos

O objetivo desta tese é a extração e a seleção de feições em imagens utilizando modelos ativos de formas adequados para uma classificação de objetos.

A proposta de extração e seleção de feições usando ASM desta tese é aplicada no

reconhecimento de faces e no monitoramento visual de painéis.

## **1.4 Contribuições**

Diferentes versões ao longo do tempo têm melhorado o ASM, mas ainda existem margens do erro de localização dos marcadores durante a comparação. Estes erros têm limitado a utilização do ASM para tarefas de reconhecimento de objetos. Nesta tese contribui-se principalmente com o tratamento das incertezas da localização e à atribuição de pesos nos marcadores, visando abordar um ASM adequado para reconhecimento de padrões. Para o ASM com aplicações em reconhecimento de faces, as incertezas de localização dos marcadores foram modeladas com Gaussianas, e foram atribuídos pesos às feições que mais discriminam às classes que representam a face das pessoas utilizando critérios com base na informação mútua ajustada (AMI) (VINH; EPPS; BAILEY, 2010), seguindo o objetivo de melhorar a taxa de reconhecimento de faces (BEHAINE; SCHARCANSKI, 2012a). Outra proposta com base em algoritmos genéticos (GA), pode melhorar de forma meta-heurística o desempenho do classificador utilizado em (BEHAINE; SCHARCANSKI, 2012a), uma vez que existe uma quantidade grande de combinações para a seleção dos marcadores que melhora a taxa de acerto (VIGNOLO et al., 2012). Outra aplicação desta tese, sugere que os marcadores adequados para fazer monitoramento visual automático de painéis analógicos são selecionados utilizando a teoria da AMI (BEHAINE; SCHARCANSKI, 2012b).

## **1.5 Estrutura do Documento**

Este documento de tese apresenta cinco capítulos: no primeiro capítulo introduz-se a tese; no capítulo dois trata-se da revisão bibliográfica, da teoria dos modelos ativos de formas e da seleção de feições; no capítulo três apresenta-se os métodos propostos nos modelos ativos de formas, que correspondem às contribuições na melhora do ASM para o reconhecimento de objetos e modificações do ASM na melhora da comparação de objetos em imagens a cores; no capítulo quatro detalha-se os materiais utilizados nas aplicações dos métodos propostos e os resultados experimentais obtidos. No final do capítulo quatro

apresenta-se a discussão. No capítulo cinco apresentam-se as conclusões a partir dos resultados obtidos e as discussões.

## 2 MODELOS ATIVOS DE FORMAS

Os modelos tradicionais de estudo no caso de objetos deformáveis são: os modelos de contornos ativos ou "*snakes*" (KASS; WITKIN; TERZOPOULOS, 1988), os modelos de *templates* deformáveis ou "*handcrafted*" (LIPSON et al., 1990) e os modelos ativos de formas ou ASM (COOTES et al., 1995). Caso os modelos abordem informação da forma e informação da textura, estes são chamados modelos de aparências ativas (AAM) (COOTES; EDWARDS; TAYLOR, 1998). O AAM por utilizar características de forma e aparência em textura (duas dimensões 2D), ficam fora do escopo desta tese.

### 2.1 Revisão Bibliográfica

Os modelos de contornos ativos ou *snakes* são curvas definidas no domínio espacial da imagem, que podem ser deslocadas por forças definidas na própria curva ou por forças definidas externamente (KASS; WITKIN; TERZOPOULOS, 1988). Os deslocamentos por forças internas mantém o modelo suave durante a deformação e os deslocamentos por forças externas fazem o modelo movimentar-se na direção das bordas do objeto ou outras feições da imagem. As *snakes* são ativas no sentido que minimizam a função de energia apresentando um comportamento dinâmico. As versões melhoradas das *snakes* procuram uma convergência dinâmica através de filtros de Kalman (TERZOPOULOS; SZELISKI, 1992), através de técnicas de condensação modeladas para ruído com distribuição não Gaussiana (ISARD; BLAKE, 1998), através de um estudo de contorno gráfico modelado com cortes manuais ou *intelligent scissors* (MORTENSEN; BARRETT, 1995), através de cortes de projeções probabilísticas (PEREZ; BLAKE; GANGNET, 2001), ou através

de mínimos flexíveis (locais ou globais) (LI; YEZZI, 2007). Estes modelos precisam de condições iniciais próximas do contorno desejado do objeto, por isso a inicialização deste tipo de modelo é crítica e é uma de suas limitações.

Os modelos de *templates* deformáveis ou *handcrafted* são paramétricos com poucos graus de liberdade e se ajustam à imagem. Estes permitem a utilização de formas complexas do objeto a detectar com funções de custo parecidas às *snakes*, é dizer minimizando a energia externa. A energia interna pode ser utilizada em segundo plano para favorecer algumas formas do objeto. As primeiras abordagens foram utilizadas para aplicações de extração de feições em imagens médicas (LIPSON et al., 1990), a extração de feições faciais (YUILLE; HALLINAN; COHEN, 1992), e a análise de imagens de ultrassom (ESCOLANO et al., 1997). A desvantagem deste tipo de abordagem é que são muito dependentes da aplicação projetada, por isso é difícil garantir que o modelo e a função custo sejam as mais adequadas para a detecção do objeto.

Os modelos ativos de formas ou ASM são modelos estatísticos de deformações de objetos, com o objetivo de realizar uma comparação ótima com a forma do objeto deformado representado na imagem. O ASM foi introduzido para resolver problemas de reconhecimento de objetos com ruído, confusos e oclusos em diferentes tipos de imagens, (por exemplo objetos elétricos, biológicos) (COOTES et al., 1995). A vantagem principal que tem o ASM quando comparado com os modelos do tipo *snake* e *handcrafted*, é que o modelo das deformações seguem parâmetros restritos às deformações reais do objeto durante o treinamento. A marcação automática de marcadores ainda não supera a marcação manual durante o treinamento (em termos do erro de localização) para alguns tipos de objetos (HILL; TAYLOR; BRETT, 2000). Outras versões de ASM têm sido propostas para melhorar a confiabilidade do processo de comparação dos marcadores, especialmente para aplicações tais como a segmentação do cérebro em imagens de ressonância magnética (DUTA; SONKA, 1998), o alinhamento de face utilizando perfis de textura (WANG et al., 2002), a segmentação de objetos em imagens médicas utilizando técnicas não lineares (VAN GINNEKEN et al., 2002), a extração de feições da face utilizando algoritmos genéticos (GA) (WAN; LAM; NG, 2005), a extração de textura da



face utilizando busca de perfis em duas dimensões (ZHENG et al., 2008), a representação esparsa do ASM em faces (LI; FENG, 2010), a extensão completa da análise de multiresolução para ASM (CERROLAZA et al., 2012). Contudo, os erros de localização dos marcadores durante a comparação ainda persistem. Existe uma versão estendida do ASM em três dimensões que utiliza descritores otimizados para melhorar a comparação dos marcadores na superfície do objeto modelado (DAVIES et al., 2002). A abordagem do ASM em três dimensões (3D) fica fora do escopo desta tese.

A maioria das abordagens ASM propostas, como as descritas em (WANG et al., 2002; VAN GINNEKEN et al., 2002; WAN; LAM; NG, 2005; ZHENG et al., 2008; LI; FENG, 2010) conta com imagens em tons de cinza (i.e. foi projetado para imagens em níveis de cinza). Mais recentemente em (TEDÁN; BECERRA; DURO, 2013) propõe-se utilizar um classificador que trata os perfis nos três canais que representam as cores (modelo de cor *Red Green Blue* RGB) separadamente, mas dita abordagem é ainda heurística.

## 2.2 Teoria dos ASMs

O ASM é construído com pontos de interesse ou marcadores. Um conjunto de marcadores que representa o objeto é chamado modelo de pontos distribuídos (PDM). Os marcadores são marcados manualmente durante o treinamento com o intuito de ter uma referência estatística robusta e confiável. Um conjunto de  $N$  marcadores representa a forma de um objeto contida em uma  $j$ -ésima imagem de treinamento como a seguir:

$$\mathbf{x}_j = [x_{j1}, y_{j1}, \dots, x_{jN}, y_{jN}]^T, \quad (1)$$

onde  $(x_{ji}, y_{ji})$  são as coordenadas do marcador  $p_i$  do PDM em uma  $j$ -ésima imagem de treinamento (COOTES et al., 1995).

Um ASM é treinado com um conjunto de imagens de um objeto dado (o conjunto de treinamento que representa os modos de variação da forma do objeto). Em um conjunto de treinamento com  $J$  imagens, tem-se  $J$  vetores na matrix de treinamento  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_J]$ . Uma análise de componentes principais (PCA) é utilizada para achar os principais modos

de deformação com base nos dados de treinamento (COOTES et al., 1995):

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{H}\mathbf{b} \quad (2)$$

onde  $\mathbf{x}$  é o modelo das deformações,  $\bar{\mathbf{x}}$  é a média de  $\mathbf{X}$ , calculada com as  $J$  imagens de treinamento como  $\bar{\mathbf{x}} = (1/J) \sum_{j=1}^J \mathbf{x}_j$ ,  $\mathbf{H}$  contém  $\varsigma$  autovetores da matriz de covariância unitária e ortogonal de  $\mathbf{X}$ , (i.e.  $(1/J)\mathbf{X}\mathbf{X}^\top$  no caso  $J \geq 2N$  ou  $(1/J)\mathbf{X}^\top\mathbf{X}$  no caso  $J < 2N$ ) (COOTES et al., 1995), associada aos autovalores  $\lambda_\ell$ , ( $\ell = 1, \dots, \varsigma$ ), então  $\mathbf{H} \in \mathbb{R}^{2N \times \varsigma}$ . O vetor  $\mathbf{b}$  de tamanho  $\varsigma \times 1$ , é expresso como  $\mathbf{b} = \mathbf{H}^\top(\mathbf{x} - \bar{\mathbf{x}})$  e define um conjunto de parâmetros de um modelo ativo de forma, onde as deformações seguem as direções ortogonais dos marcadores (COOTES et al., 1992, 1995). Então, a forma de  $\mathbf{x}$  pode ser ajustada pelas mudanças dos elementos em  $\mathbf{b}$ . A consistência da forma é fornecida pelas restrições das variações do  $\ell$ -ésimo parâmetro  $b_\ell$ , a saber,  $\pm 3\sqrt{\lambda_\ell}$  (COOTES et al., 1995).

Também é criada uma estrutura de aparência dos níveis de cinza ou perfis. A seguir, apresenta-se alguns conceitos relevantes chamados vetores do perfil, vetores derivadas do perfil, e vetores derivada do perfil normalizado. Estes conceitos são aplicados às imagens de tons de cinza do objeto para a busca de marcadores em tais imagens. Uma estrutura do perfil é amostrada com  $\epsilon$  *pixels* dos níveis de cinza nas normais às bordas centradas em cada marcador (COOTES; TAYLOR, 1993). A Figura 3 ilustra a localização das normais nos  $N = 12$  marcadores de um PDM de um objeto.

O vetor do perfil no  $i$ -ésimo marcador de uma  $j$ -ésima imagem de treinamento é:

$$\mathbf{f}_{ji_\ell} = [f(1)_{ji_\ell}, \dots, f(\epsilon)_{ji_\ell}], \quad (3)$$

onde  $\mathbf{f}_{ji_\ell}$  é o vetor do perfil, e  $f(\cdot)_{ji_\ell}$  é o nível de cinza do *pixel* em termos de luminância ( $\ell$ ) capturados ao longo da normal no contorno local do  $i$ -ésimo marcador (COOTES; TAYLOR, 1993; WAN; LAM; NG, 2005). O vetor derivada do perfil no  $i$ -ésimo marcador, é expresso pelas mudanças dos níveis de cinza que ocorrem entre *pixels* adjacentes no vetor do perfil indicado na Equação (3):

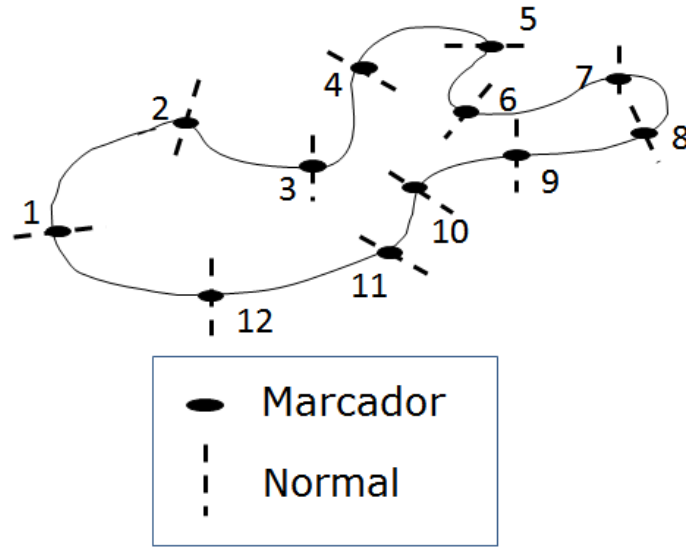


Figura 3: Ilustração da localização das normais nos  $N = 12$  marcadores de um PDM de um objeto.

$$\mathbf{df}_{ji_\epsilon} = [f(2)_{ji_\epsilon} - f(1)_{jq_\epsilon}, \dots, f(\epsilon)_{ji_\epsilon} - f(\epsilon - 1)_{ji_\epsilon}]^\top, \quad (4)$$

onde  $\mathbf{df}_{ji_\epsilon}$  denota o vetor derivada do perfil. Os resultados experimentais mostram que o vetor derivada do perfil normalizado tem melhor desempenho em termos de localização do marcador quando comparado com o vetor do perfil, principalmente porque o vetor derivada do perfil normalizado é mais robusto às mudanças de luminosidade local (COOTES; TAYLOR, 1993). O vetor derivada do perfil normalizado é definido como:

$$\mathbf{g}_{ji_\epsilon} = \frac{\mathbf{df}_{ji_\epsilon}}{\sum_{b=1}^{(\epsilon-1)} |df^b(b)_{ji_\epsilon}|}, \quad (5)$$

onde  $\mathbf{g}_{ji_\epsilon}$  denota o vetor derivada do perfil normalizado,  $|df^b(b)_{ji_\epsilon}|$  é o valor absoluto do  $b$ -ésimo elemento do vetor derivada do perfil associado ao  $i$ -ésimo marcador.

O modelo de aparência do objeto é treinado com  $J$  imagens (com os marcadores colocados com antecedência). Um conjunto de  $J$  vetores derivada do perfil normalizado (um por cada imagem de treinamento para o  $i$ -ésimo marcador) estão dispostos em uma

matriz, em que as colunas são os  $J$  vetores derivadas do perfil normalizado:

$$\mathbf{G}_{i_\varepsilon} = [\mathbf{g}_{1i_\varepsilon}, \mathbf{g}_{2i_\varepsilon}, \dots, \mathbf{g}_{Ji_\varepsilon}], \quad (6)$$

onde  $\mathbf{G}_{i_\varepsilon}$  é a matriz chamada simplesmente de matriz de perfis de luminância no  $i$ -ésimo marcador.

A estimativa da posição de cada  $i$ -ésimo marcador na imagem de teste apóia-se no procedimento de comparação entre perfis ao longo da normal. A função a ser minimizada durante a comparação dos perfis é a distância de Mahalanobis definida a seguir (COOTES, 2000):

$$M_d(\mathbf{g}_{i_\varepsilon}^t) = (\mathbf{g}_{i_\varepsilon}^t - \bar{\mathbf{g}}_{i_\varepsilon})^\top \mathbf{G}_c^{-1} (\mathbf{g}_{i_\varepsilon}^t - \bar{\mathbf{g}}_{i_\varepsilon}), \quad (7)$$

onde  $\mathbf{g}_{i_\varepsilon}^t$  é o vetor derivada do perfil normalizado da  $t$ -ésima imagem de teste no  $i$ -ésimo marcador,  $M_d(\mathbf{g}_{i_\varepsilon}^t)$  é a distância de Mahalanobis entre  $\mathbf{g}_{i_\varepsilon}^t$  e o modelo de aparência. Assume-se que o modelo de aparência tem uma estatística Gaussiana, onde a média é calculada  $\bar{\mathbf{g}}_{i_\varepsilon} = (1/J) \sum_{j=1}^J \mathbf{g}_{ji_\varepsilon}$  e  $\mathbf{G}_c$  representa a matriz de covariância de  $\mathbf{G}_{i_\varepsilon}$ , (i.e.  $\mathbf{G}_c = (1/J) \mathbf{G}_{i_\varepsilon} \mathbf{G}_{i_\varepsilon}^\top$ ). Uma vez obtida a minimização da Equação (7) ao longo da normal é possível obter uma estimativa da posição de cada marcador candidato da  $t$ -ésima imagem de teste. O procedimento de comparação da forma no ASM procura a pose e os parâmetros da forma que encaixem melhor em uma instância do modelo  $\mathbf{x}$ , até um novo conjunto de pontos  $I$  da  $t$ -ésima imagem de teste. Este objetivo é atingido iterativamente minimizando a soma das distâncias quadráticas entre o modelo correspondente e os marcadores sob a imagem  $|I - T_{\tau, \xi, \theta}(\bar{\mathbf{x}} + \mathbf{H}\mathbf{b})|^2$ , onde  $T_{\tau, \xi, \theta}(\bar{\mathbf{x}} + \mathbf{H}\mathbf{b})$  denota a posição do PDM na imagem fazendo uma rotação por  $\theta$ , uma escala por  $\xi$  e deslocado por  $(x, y)_\tau$  (COOTES, 2000).

### 2.3 Seleção de Feições

As feições são atributos que referem-se às características dos dados (LADHA; DE-EPA, 2011). Usualmente as feições são especificadas ou selecionadas antes de serem

coletadas. As feições podem ser contínuas ou discretas. Em geral, as feições podem ser como a seguir:

- a) Feições relevantes: são as feições que influenciam na saída de um sistema de reconhecimento de padrões e não podem ser substituídas por outras feições;
- b) Feições irrelevantes: são as feições que não influenciam na saída de um sistema de reconhecimento de padrões e cujos valores são gerados aleatoriamente para cada amostra na entrada;
- c) Feições redundantes: são as feições que podem assumir a mesma função de outras feições.

Exemplos de feições presentes objetos podem ser, no caso das faces: crominância  $C_b$  e crominância  $C_r$  do íris, crominância  $C_b$  e crominância  $C_r$  no limiar do queixo, etc; e no caso dos painéis: posição em  $x$  e posição em  $y$  no ponteiro, etc. As feições relevantes, as feições irrelevantes e as feições redundantes são difíceis de reconhecer na presença de ruído. Quando o sistema de reconhecimento de padrões pretende classificar as feições coletadas sem ser selecionadas com antecedência, é possível que o ruído em todas as feições impacte negativamente o desempenho do sistema em termos da taxa de classificação (DUDA; HART; STORK, 2001). Caso o conjunto de feições não seja selecionado ou reduzido no pré-processamento, o desempenho do classificador em termos do tempo de execução também pode comprometer os resultados em tempo real do sistema de reconhecimento.

O termo seleção de feições refere-se às técnicas que selecionam as melhores feições (feições relevantes) dentre o conjunto de feições na entrada do sistema de reconhecimento, e o termo extração de feições se refere aos métodos que criam novas feições a partir de transformações ou combinações das feições selecionadas na entrada (JAIN; DUIN; MAO, 2000). As feições selecionadas têm um significado físico do processo ou do objeto a ser reconhecido, entretanto as feições extraídas podem não ter um significado físico claro. A escolha entre a seleção de feições e a extração de feições depende do tipo de aplicação e dos dados disponíveis durante o treinamento. A seleção de feições é importante porque

permite as seguintes vantagens (LADHA; DEEPA, 2011): a limitação do requisito de armazenamento dos dados; a diminuição dos tempos de execução de processos; a melhora da qualidade dos dados; a remoção da redundância, da irrelevância e do ruído; o incremento da exatidão do sistema de reconhecimento e a melhora na forma da visualização dos dados.

O objetivo das técnicas de seleção de feições é identificar um conjunto de  $R_{\mathfrak{F}}$  feições dado um conjunto inicial de  $N_{\mathfrak{F}}$  feições com o menor erro de classificação, onde  $R_{\mathfrak{F}} \leq N_{\mathfrak{F}}$ . Se denotarmos a  $\mathfrak{F}$  como a função critério que seleciona o conjunto ótimo de feições, então uma função critério é

$$\mathfrak{F} = 1 - \wp, \quad (8)$$

onde  $\wp$  é a probabilidade do erro da classificação. A Equação (8) infere que teria-se que examinar o erro de classificação das combinações possíveis na entrada e selecionar o conjunto de feições que maximiza o valor da função critério  $\mathfrak{F}$ . Entretanto, no tratamento de imagens nem sempre é possível ter uma quantidade suficiente de dados no conjunto de treinamento para ter uma estatística robusta ao examinar a probabilidade de erro da classificação.

Existem três tipos de métodos fundamentais para a seleção de feições, cuja classificação pode ser dada como a seguir:

- a) Métodos de filtragem: são os métodos que utilizam uma função de discriminação e não dependem do classificador no processo de aprendizagem. A função de discriminação pode ser baseada na mínima redundância e máxima relevância, no ganho máximo de informação, na mínima correlação, na mínima informação mútua, na distancia máxima de separação, etc;
- b) Método empacotados ou *wrappers*: são os métodos que utilizam uma hipótese nos dados de entrada que é testada pelo classificador na etapa de aprendizado. O objetivo dos *wrappers* é obter uma função de predição da classificação. A função de predição do classificador pode ser apoiada pelos algoritmos genéticos (GA) ou pelo aprendizado

de máquina, como por exemplo máquinas de vetores de suporte (SVM);

- c) Métodos híbridos: são os métodos que misturam recursos dos métodos de filtragem com os recursos dos *wrappers*. Os métodos híbridos também podem utilizar procedimentos heurísticos.

A vantagem dos *wrappers* quando comparados com os métodos de filtragem é que em geral o seu desempenho em termos da exatidão é maior. A desvantagem é que consomem mais recursos computacionais e dependem do classificador.

A forma como a seleção de feições é implementada impacta no desempenho do sistema de reconhecimento, tanto em termos de exatidão quanto em termos de recursos computacionais e tempo de execução. Os algoritmos conhecidos para a seleção de feições são: busca exaustiva, melhor feição individual, *Sequential Forward Selection* (SFS) (HUANG et al., 2010), *Sequential Backward Selection* (SBS) (HUANG et al., 2010), *Sequential Backward Floating Search* (SBFS) (PUDIL; NOVOVICOVÁ; KITTLER, 1994), *Sequential Forward Floating Search* (SFFS) (JAIN; ZONGKER, 1997). A seleção de feições pode ser apoiada através de outras técnicas e algoritmos adequados, a saber, informação mútua (MI) (PENG; LONG; DING, 2005) e algoritmos genéticos (GA) (LEI, 2012). Na avaliação dos algoritmos para seleção de feições existe um compromisso entre recursos computacionais e o desempenho em termos da taxa de classificação (DASH; LIU, 1997; MOLINA; BELANCHE; NEBOT, 2002; LADHA; DEEPA, 2011).

### **3 INOVAÇÕES PROPOSTAS NOS MODELOS ATIVOS DE FORMAS**

Neste capítulo são apresentadas as inovações propostas nos modelos ativos de formas. As ditas inovações correspondem às contribuições na melhora do ASM para o reconhecimento de objetos apresentada na seção 3.1 e às modificações no ASM para a melhora da comparação de objetos em imagens a cores apresentada na seção 3.2.

O ASM proposto para reconhecimento de objetos baseia-se em um esquema de seleção de marcadores em ASM para melhorar o desempenho em aplicações de reconhecimento de objetos. O esquema proposto seleciona pontos marcadores consistentes onde feições relevantes do objeto estão localizadas, e associa pesos altos às feições correspondentes encontradas na etapa de classificação dos objetos. Os pesos no esquema de seleção dos marcadores indicam o nível de discriminação das feições em uma localização dentre as diferentes feições localizadas para todas as classes do objeto que se pretende reconhecer. As classes do objeto são modeladas com funções Gaussianas cujos parâmetros são extraídos adequadamente para tratar as incertezas estimadas da localização dos marcadores durante o treinamento do método de reconhecimento baseado no ASM. Os resultados são promissores e indicam que o método proposto procura melhorar o desempenho do ASM, levando a melhoras dos resultados finais de reconhecimento de objetos em imagens tais como faces e painéis, cujos resultados experimentais estão ilustrados na seção 4.1 e na seção 4.2 respectivamente.

O ASM modificado para a melhora da comparação de objetos em imagens a cores, baseia-se na definição adequada de conjuntos de marcadores que representam o objeto e



na extensão de perfis de tons de cinza para os diferentes canais que representam as cores. Particiona-se o conjunto de marcadores em dois conjuntos de pontos, a saber, o conjunto de marcadores identificados na imagem, e o conjunto de marcadores que são interpolados a partir dos marcadores identificados. Uma vez definido o conjunto de marcadores que irão ser utilizados no PDM para representar o objeto, um perfil *Red Green Blue* (RGB) (modelo de cor RGB) é detectado em torno de cada um destes marcadores. Utiliza-se um perfil RGB na vizinhança de cada uma dos marcadores detectados sobre uma imagem de entrada (i.e. imagem a cores representada em RGB) do objeto, para melhorar a comparação da forma. Um modelo de aparência local é criado a partir destes perfis de tons de cinza durante a etapa de treinamento de um ASM, que por sua vez, é também utilizado nas imagens de entrada do objeto. O perfil de tons de cinza auxilia na detecção dos marcadores através da estimativa das localizações mais parecidas da imagem de entrada. Os resultados experimentais sugerem que o método proposto melhora o desempenho do ASM em objetos representados em imagens a cores, levando à melhoras dos resultados finais de leitura no monitoramento visual automático de multipainéis quando contrastado com a leitura humana e outras abordagens do ASM, cujos resultados experimentais estão ilustrados na seção 4.3.

### **3.1 Modelos Ativos de Formas e Reconhecimento de Objetos**

Apresenta-se nesta seção um método baseado na informação mútua ajustada (AMI) (VINH; EPPS; BAILEY, 2010) para a seleção de feições em objetos com o intuito de adequar o ASM em reconhecimento de padrões deformáveis.

As feições de um objeto são características importantes para o seu reconhecimento e podem ser adquiridas sem contato através de sensores de imagens. Nos métodos de modelagem e reconhecimento de objetos, tais como o ASM, um número ( $N$ ) de feições são capturadas em uma imagem do objeto usada como entrada, mas somente  $R$  ( $R < N$ ) feições do objeto são úteis para a caracterização da face, e as outras ( $N - R$ ) feições tem pequenas contribuições ou são ruidosas. Essa questão tem sido negligenciada na estrutura do ASM adotado na comparação de objetos. Os erros da comparação do objeto com

ASM podem ser altos em algumas localidades do objeto (HILL; COOTES; TAYLOR, 1996), apesar de melhoras recentes feitas na técnica do ASM (WAN; LAM; NG, 2005), (JUNMO; MüJDAT; ALAN, 2007). Muito embora implementações recentes do ASM tenham melhorado a exatidão da localização dos marcadores em imagens de objetos (por exemplo no caso de faces (ZHENG et al., 2008)), a detecção de feições do objeto com variações de pose e iluminação é ainda desafiadora (MILBORROW; NICOLLS, 2008). A detecção inconsistente de feições do objeto acontece frequentemente porque o ASM treinado converge em direção às bordas salientes da imagem, e se estas bordas salientes forem ruidosas ou distorcidas pelo sombreamento ou iluminação, pode acontecer comparações erradas de feições. Contudo, feições do objeto que são localizadas consistentemente no conjunto de treinamento são mais relevantes para a caracterização e a discriminação do problema de classificação do objeto.

O reconhecimento de objetos em imagens pode ser entendido como a classificação de um objeto dentre  $K$  classes de objetos, assumindo que a forma, pose e iluminação podem mudar levemente com relação aos dados de treinamento. No ASM, a forma dos objetos é modelada por modelos de pontos distribuídos (PDMs), como é ilustrado na Figura 4(a) no caso de faces. Os pontos de um PDM são chamados marcadores, e a localização desses marcadores nas imagens dos objetos podem ter erros de localização ou erros de comparação (veja-se a Figura 4(b) no caso de faces). Um ASM para reconhecimento de objetos é treinado com base em um conjunto de imagens de treinamento, e  $N$  pontos do PDM  $S_{k,\epsilon}$  são utilizados para representar a forma de cada objeto da  $k$ -ésima classe do objeto,  $k = 1, \dots, K$ , a saber  $S_{k,\epsilon} = \{p_i(x_i + \epsilon_{x_i}, y_i + \epsilon_{y_i})\}^k$ ,  $i = 1, \dots, N$ , onde  $(x_i, y_i)$  são as coordenadas de um ponto do PDM  $p_i$  e  $(\epsilon_{x_i}, \epsilon_{y_i})$  são os erros de localização. Cada marcador  $p_i$  do PDM representa uma localização de uma feição relevante do objeto (i.e. bordas, cantos, limiar, etc.). O aspecto visual (aparência) do marcador na imagem do objeto é descrito pelo conjunto de feições (por exemplo posição no eixo  $x$ , posição no eixo  $y$ , luminância, crominância, textura, etc.), aqui denotada  $\{F_{m,i}\}$ ,  $m = 1, \dots, M$ , onde  $M$  é o número de feições da imagem utilizada por cada marcador.

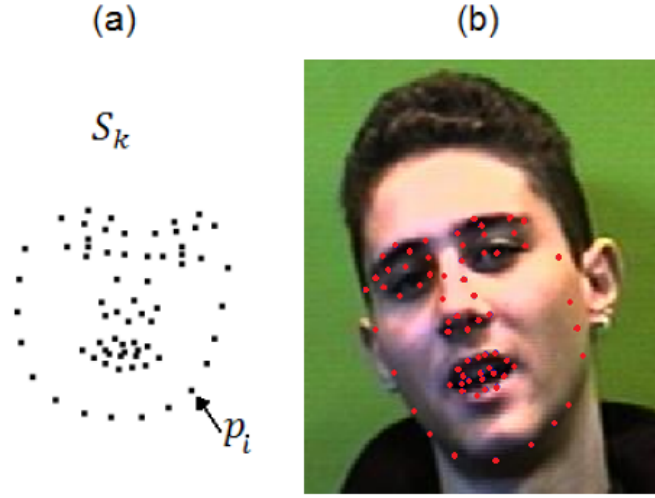


Figura 4: Ilustração de: (a) marcadores  $p_i$  (PDM) utilizado para modelar uma face no ASM; e (b) Localização de marcadores em uma imagem da face.

### 3.1.1 Composição dos Vetores de Feições com Marcadores Chave em ASMs

Nesta abordagem, cada um dos  $N$  marcadores  $p_i$  do PDM é descrito pela média  $\mu_{F_{m,i}}$  e a variância  $\sigma_{F_{m,i}}^2$  das medições na imagem da  $m$ -ésima feição em todas as  $J$  amostras das imagens de treinamento em uma classe do objeto ( $F_{m,i}^j, j = 1, \dots, J$ )

$$\mu_{F_{m,i}} = \frac{1}{w^2} \sum_{\hat{r}=1}^w \sum_{\hat{q}=1}^w \mu_{m,i}(\hat{r}, \hat{q}) \quad (9)$$

$$\sigma_{F_{m,i}}^2 = \max_{\hat{r}, \hat{q} \in W} \{ \sigma_{m,i}^2(\hat{r}, \hat{q}) \} \quad (10)$$

com

$$\mu_{m,i}(\hat{r}, \hat{q}) = (1/J) \sum_{j=1}^J F_{m,i}^j(\hat{r}, \hat{q}) \quad (11)$$

e

$$\sigma_{m,i}^2(\hat{r}, \hat{q}) = (1/J) \sum_{j=1}^J (F_{m,i}^j(\hat{r}, \hat{q}) - \mu_{m,i}(\hat{r}, \hat{q}))^2, \quad (12)$$

onde  $(\hat{r}, \hat{q})$  são as coordenadas dos *pixels* contidos na janela  $W = w \times w$  centrada no  $i$ -ésimo marcador  $p_i$  ASM, e calculada para a  $m$ -ésima feição da imagem. A variância

máxima da janela foi selecionada na Equação (10) para levar em conta a variabilidade da feição da vizinhança em  $w \times w$  do  $i$ -ésimo marcador  $p_i$  ASM. Os erros de localização são assumidos com estatística Gaussiana de média zero e variância  $\sigma_i^2$  para cada marcador  $p_i$  durante o treinamento ASM (SHI; SAMAL; MARX, 2006), e adota-se  $w = 2 \max \{\sigma_\epsilon\}$ , onde  $\sigma_\epsilon$  é o desvio padrão dos erros de localização de todos os marcadores, medidos durante o treinamento do ASM. Cada  $k$ -ésima classe do objeto pode ser descrita pela média  $\mu_{F_{m,i}}^k$  e a variância  $\sigma_{F_{m,i}}^2$  de cada um dos  $N$  marcadores  $p_i$  do PDM, utilizando-se a Equação (9) e a Equação (10) para cada classe, respectivamente. Uma  $m$ -ésima feição no  $i$ -ésimo marcador para todas as  $K$  classes do objeto pode ser representada compactamente

$$\mathbf{u}_{F_{m,i}} = [\mu_{F_{m,i}}^1, \dots, \mu_{F_{m,i}}^K]^\top \quad (13)$$

onde  $\mathbf{u}_{F_{m,i}}$  é o vetor da  $m$ -ésima feição no  $i$ -ésimo marcador das classes, aqui também chamado vetor de feições.

### 3.1.2 Seleção de Feições Consistentes para Reconhecimento de Objetos com ASMs

O desempenho da localização dos marcadores no ASM tende a melhorar com conjuntos numerosos de marcadores (MILBORROW; NICOLLS, 2008). Entretanto, o desempenho da classificação de objetos baseado na localização dos marcadores do ASM diminui quando a dimensão do problema aumenta. Descreve-se nesta seção a abordagem da seleção de  $R$  marcadores que estão consistentemente localizados durante o treinamento do ASM ( $R \leq N$ ), conseguindo-se uma melhor separação entre classes de objetos.

Deseja-se ter uma expressão adequada que avalie o grau de discriminação entre classes, i.e. uma relação inversa da informação mútua entre as diferentes feições coletadas em cada marcador  $p_i$ . A AMI é uma medida adimensional normalizada da dependência mútua entre duas variáveis e é adequada para tratamento de dados aglomerados ou *clusters* (VINH; EPPS; BAILEY, 2010), cujo cálculo é descrito no apêndice A. Seja  $A(u, v) = (MI(u, v) - E\{MI(u, v)\}) / (\max\{\hat{H}(u), \hat{H}(v)\} - E\{MI(u, v)\})$  a AMI entre as variáveis  $u$  e  $v$ ,  $0 \leq AMI(u, v) \leq 1$ , onde  $MI(u, v)$ ,  $E\{MI(u, v)\}$  denota a informação mútua entre as variáveis  $u$  e  $v$ , e o seu valor esperado, respectivamente;  $\hat{H}(u)$

e  $\hat{H}(v)$  são as entropias do  $u$  e  $v$ . Pesa-se a consistência  $a_i$  de um marcador  $p_i$  no conjunto de treinamento como a seguir:

$$a_i = \frac{1}{M} \sum_{m=1}^M \left( \frac{1}{\sum_{n=1}^N A(\mathbf{u}_{F_{m,n}}, \mathbf{u}_{F_{m,i}})} \right), \quad (14)$$

onde  $A(\mathbf{u}_{F_{m,n}}, \mathbf{u}_{F_{m,i}})$  é calculada entre os vetores de feições  $\mathbf{u}_{F_{m,n}}$  e  $\mathbf{u}_{F_{m,i}}$  considerando todas as amostras dos marcadores  $p_i$  no conjunto de treinamento (englobando todas as classes do objeto), e para os  $N$  marcadores  $p_n$ , (i.e.  $\mathbf{u}_{F_{m,i}} \in \mathbb{R}^{K \times 1}$  tem  $K$  valores, um  $\mu_{F_{m,i}}^k$  para cada  $k$ -ésima classe de face como ilustrado na Equação (13)). Valores altos de  $a_i$  refletem baixa AMI no conjunto de treinamento entre os marcadores  $p_i$  e os outros marcadores  $p_n$ , ou menor interdependência. O valor do  $R$  depende do problema (TURK; PENTLAND, 1991). Seleciona-se os  $R$  marcadores com os maiores valores  $a_i$  ( $R \leq N$ ) como o conjunto de marcadores menos interdependente, reduzindo-se a dimensionalidade do problema de reconhecimento de objetos. De fato,  $a_i$  é interpretado como uma distinção das medições extraídas no  $p_i$ . Em um PDM de um ASM padrão os marcadores  $p_i$  não possuem pesos que discriminem as diferentes classes de um objeto. Neste método proposto um PDM pode adquirir pesos discriminantes entre classes. A Figura 5 (a) ilustra os pesos discriminantes  $a_i$  nos  $N = 12$  marcadores de um PDM de um objeto. Uma vez calculados os pesos  $a_i$  a partir das localizações fornecidas por um ASM, então é possível selecionar os  $R$  marcadores mais discriminantes de um PDM. A Figura 5 (b) ilustra a seleção dos  $R = 5$  marcadores (em vermelho) mais discriminantes de um PDM com  $N = 12$  marcadores para a classificação de um objeto.

Dada uma imagem do objeto, tem-se  $N$  marcadores  $p_i$  localizados, fornecidos pelo treinamento ASM onde as feições do objeto da imagem foram medidas. A probabilidade dos valores assumidos pela feição  $m$  da imagem no marcador  $p_i$  no conjunto de treinamento de uma classe objeto  $k$  são dados por uma Gaussiana  $G_{m,i}^k = G(\mu_{F_{m,i}}^k, \sigma_{F_{m,i}}^2{}^k)$ , onde  $\mu_{F_{m,i}}^k$  e  $\sigma_{F_{m,i}}^2{}^k$  são a sua media e variância, conforme Equações (9) e (10) respectivamente. Então, considerando todos os  $N$  marcadores e as  $M$  feições da imagem do objeto medidas na localização dos marcadores, uma mistura de Gaussianas é utilizada para mo-

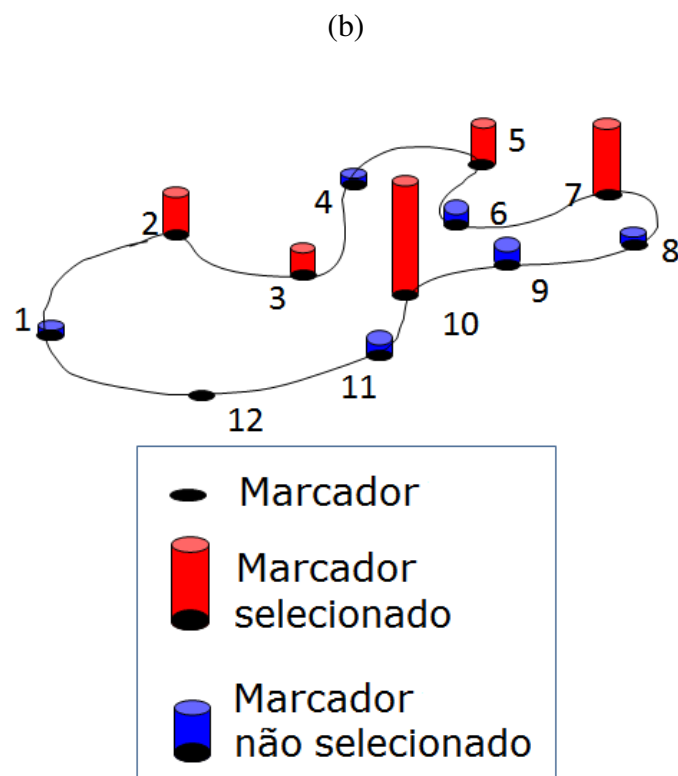
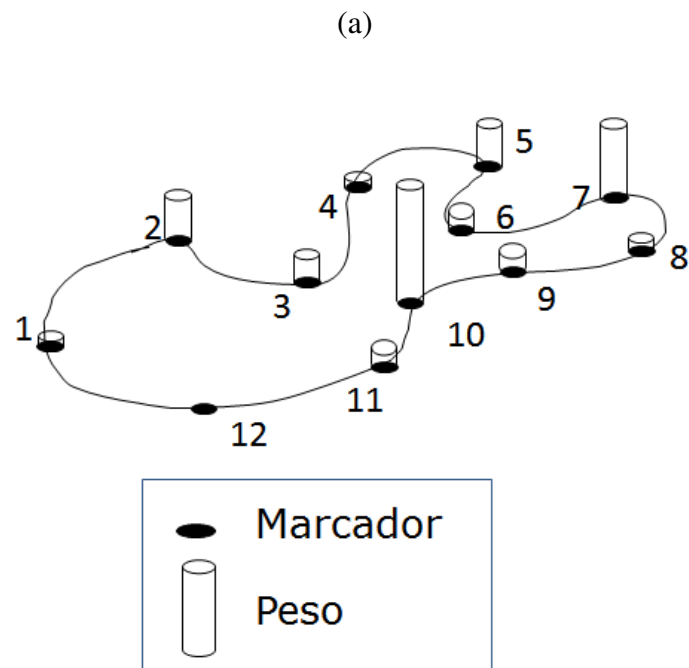


Figura 5: Representação de um PDM de 12 marcadores com a (a) Ilustração dos pesos discriminantes  $a_i$  nos marcadores, e a (b) Ilustração da seleção de  $R = 5$  marcadores (em vermelho) para a classificação de um objeto.

delar estatisticamente a classe de objeto  $k$  como  $\Omega_k$ , onde  $\Omega_k = \sum_{i=1}^N \sum_{m=1}^M d_{X_{m,i}^k} G_{m,i}^k$ , e  $d_{X_{m,i}^k}$  é o peso da componente da mistura de Gaussianas  $G_{m,i}^k$ . Detalhes de uma estimativa de misturas de Gaussianas são dados em (FIGUEIREDO; JAIN, 2002).

A localização de alguns marcadores pode mudar substancialmente em diferentes imagens do objeto, gerando-se comparações não confiáveis das feições nas localizações dadas. Portanto,  $\Omega_k$  é mudado nesta abordagem para levar em consideração a confiabilidade dos marcadores no processo de comparação das feições do objeto. Neste método de comparação das feições do objeto, designa-se pesos mais altos aos pontos mais consistentes (i.e. a um marcador  $p_i$  pode ser atribuído um peso diferentemente para cada  $k$  classe de objeto, refletindo a relevância do marcador para a caracterização de feições do objeto naquela localização). Portanto, designando-se um peso diferente para cada marcador  $p_i$  em cada  $k$  classe do objeto, a confiabilidade da medição extraída no  $p_i$  é considerada durante a comparação das feições do objeto na classe  $k$ , e o  $\Omega_k$  mudado, a saber,  $\Omega_{ck}$  é:

$$\Omega_{ck} = \frac{1}{R} \sum_{i=1}^R a_i \left( \frac{1}{M} \sum_{m=1}^M \alpha G_{m,i}^k + (1 - \alpha) B_i^k \right) \quad (15)$$

onde  $\Omega_{ck}$  é um valor escalar,  $B_i^k$  indica a confiabilidade das feições medidas extraída no  $p_i$  na classe  $k$ , onde  $0 < B_i^k \leq 1$ ,  $0 < \alpha \leq 1$ , e  $B_i^k = (1/(JM)) \sum_{j=1}^J \sum_{m=1}^M G_{m,i}^{k,j}$ , e  $G_{m,i}^{k,j}$  representa a probabilidade segundo uma distribuição gaussiana da feição  $m$  no marcador  $p_i$  na image do objeto  $j$  da classe do objeto  $k$ . Valores altos de  $B_i^k$  sugerem alta confiabilidade das feições medidas no  $p_i$  na classe  $k$ . O termo  $(1 - \alpha) B_i^k$  balanceia a confiabilidade da localização onde as feições são medidas ( $p_i$ ) e o quanto a medida obtida da classe  $k$  é representativa. O valor do  $\alpha$  é calculado para maximizar o valor preditivo positivo ( $PPV$ ) durante o treinamento. O  $PPV = TP/(TP + FP)$ , onde o  $TP$  e  $FP$  são as taxas do verdadeiro positivo e do falso positivo calculadas com relação à referência, respectivamente (FLETCHER; FLETCHER, 2005). O  $\alpha$  ótimo foi estimado utilizando-se um algoritmo de validação cruzada nos dados de treinamento (GEISSER, 1993). O algoritmo para a estimativa do  $\alpha$  ótimo é ilustrado no apêndice B. Lembre-se que  $a_i$  pesa a qualidade da distinção das medições extraídas no marcador  $p_i$  dentre o conjunto original de  $N$  marcadores (ver a Equação (14)).

Dada uma imagem do objeto de entrada, a classe  $k'$  do objeto que maximiza  $\Omega_{ck}$  ( $\forall k \in \{K\}$  onde  $\{K\}$  é o conjunto de todas as classes do objeto) indica a classe do objeto mais provável, que tem as feições do objeto mais parecidas com a imagem do objeto de entrada. Portanto, a imagem do objeto de entrada é designada pertencer à classe do objeto  $k'$  após fazer uma busca no conjunto de todas as classes do objeto, e seleciona-se

$$k' = \arg \max_{\forall k \in \{K\}} \{\Omega_{ck}\}. \quad (16)$$

## 3.2 Modelos Ativos de Formas Modificados

Nesta seção apresenta-se um esquema ASM modificado que procura melhorar a robustez da comparação do ASM. Nesta abordagem particiona-se o conjunto de marcadores em dois conjuntos de pontos, a saber, o conjunto de marcadores identificados na imagem, e o conjunto de marcadores que são interpolados a partir dos marcadores identificados. Uma vez definido o conjunto de marcadores que irão ser utilizados no PDM para representar o objeto, um perfil de tons de cinza é detectado em torno de cada um destes marcadores. O esquema proposto utiliza um perfil de tons de cinza na vizinhança de cada uma dos marcadores detectados sobre uma imagem de entrada multicanal (i.e. imagem a cores) do objeto, para melhorar a comparação da forma. Este perfil de tons de cinza fornece um modelo de aparência local que é usado para o treinamento de um ASM, que por sua vez, é também utilizado nas imagens de entrada do objeto. O perfil de tons de cinza auxilia na detecção dos marcadores através da estimativa das localizações mais parecidas da imagem de entrada, conforme vai ser detalhado nesta seção.

### 3.2.1 Marcadores para a Modelagem das Deformações de um Objeto

Nesta seção foca-se na região de interesse (ROI) da imagem onde os objetos estão representados. As mudanças (i.e. deformações, rotações ou deslocamentos) que o objeto sofre são modeladas pelos marcadores interpolados e identificados de um PDM. O  $j$ -ésimo conjunto PDM (na  $j$ -ésima imagem de treinamento) corresponde apenas a uma parte dos  $Q$  marcadores, denotado como  $\{P_{jq}\}$ ,  $q = 1, \dots, Q$ . Nesta abordagem particiona-se o



conjunto PDM em dois conjuntos de marcadores, como explicado a seguir. Um conjunto de marcadores capta as feições importantes do objeto, como cantos e bordas, e é chamado de "marcadores identificados", e é denotado como  $\{P_{ja_n}\}$ ,  $n = 1, \dots, N$ . O segundo conjunto de marcadores é interpolado entre os marcadores identificados consecutivos do objeto, e somente representa as bordas do objeto (uma vez que normalmente estão localizados nos limites do objeto). Chama-se o segundo conjunto como "marcadores interpolados", e é denotado como  $\{P_{jb_l}\}$ ,  $l = 1, \dots, L$ . Dois marcadores identificados consecutivos podem ter marcadores interpolados entre eles. A Figura 6 ilustra os conjuntos de marcadores identificados e marcadores interpolados para o caso de um objeto que representa a forma de um ponteiro de um painel analógico. Normalmente, os marcadores identificados são localizados manualmente na fase de treinamento do ASM, mas os marcadores interpolados são localizados automaticamente durante a fase de treinamento, o que simplifica o processo já que existem menos marcadores que precisam ser localizados manualmente.

No  $j$ -ésimo conjunto PDM  $\mathbf{x}_j$  (que corresponde à  $j$ -ésima imagem do conjunto de treinamento), os dois conjuntos de marcadores representam de forma compacta a forma de um objeto, tal como indicado abaixo:

$$\mathbf{x}_j = [x_{ja_1}, y_{ja_1}, x_{jb_1}, y_{jb_1}, \dots, x_{jb_L}, y_{jb_L}, x_{ja_N}, y_{ja_N}]^T, \quad (17)$$

onde  $(x_{ja_n}, y_{ja_n})$  são as coordenadas do  $n$ -ésimo "marcador identificado",  $(x_{jb_l}, y_{jb_l})$  são as coordenadas do  $l$ -ésimo "marcador interpolado", e  $Q = N + L$ .

O ASM usado para modelar o objeto é treinado com  $J$  imagens onde o objeto aparece em localizações diferentes da imagem. Nas imagens de treinamento os marcadores identificados são localizados, e a matriz de treinamento dos PDMs  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_J]$  é obtida. Como mencionado anteriormente, a análise de componentes principais (PCA) é usada para encontrar os principais modos de deformação da forma do objeto com base nos dados de treinamento  $\mathbf{X}$ . Os modos de deformação são associados aos autovetores da matriz de covariância de  $\mathbf{X}$ , e a consistência da forma é alcançada mediante a definição de restrições aos parâmetros, em termos dos autovalores calculados. A referência das defor-

mações é a média dos PDMs treinados. Portanto, cada  $z$ -ésima instância de deformação do objeto é parametrizada como a seguir:

$$\mathbf{x}_z = \bar{\mathbf{x}} + \mathbf{H}\mathbf{m}_z, \quad (18)$$

onde  $\bar{\mathbf{x}}$  é a média de  $\mathbf{X}$ , calculada com as  $J$  imagens de treinamento como  $\bar{\mathbf{x}} = (1/J) \sum_{j=1}^J \mathbf{x}_j$ ,  $\mathbf{H}$  contém  $\varsigma$  autovetores da matriz de covariância unitária e ortogonal de  $\mathbf{X}$ , associada aos autovalores  $\lambda_\iota$ , ( $\iota = 1, \dots, \varsigma$ ), então  $\mathbf{H} \in \mathbb{R}^{2Q \times \varsigma}$ . O vetor  $\mathbf{m}_z$  de tamanho  $\varsigma \times 1$  é expresso como  $\mathbf{m}_z = \mathbf{H}^\top (\mathbf{x}_z - \bar{\mathbf{x}})$  e define um conjunto de parâmetros de um modelo ativo de forma para a  $z$ -ésima instância. A consistência da forma é fornecida com as restrições das variações do  $\iota$ -ésimo parâmetro  $m_{z_\iota}$  em  $\pm 3\sqrt{\lambda_\iota}$ . A Equação (18) representa o modelo para todas as formas no conjunto de treinamento de PDMs  $\mathbf{X}$ , (i.e. um  $z$ -ésima instância representa a  $j$ -ésima forma  $\mathbf{x}_z = \mathbf{x}_j$ , é calculada com os parâmetros  $\mathbf{m}_z = \mathbf{H}^\top (\mathbf{x}_j - \bar{\mathbf{x}})$ ).

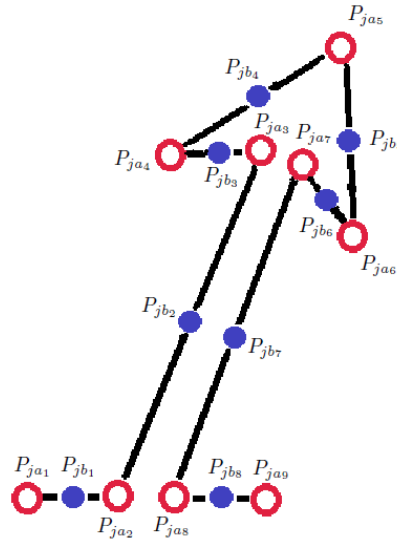


Figura 6: Ilustração de um PDM utilizando marcadores identificados  $P_{a_n}$  (círculos vermelhos vazios), e os marcadores interpolados (círculos azuis) que modelam um ponteiro de um painel analógico utilizando um ASM com  $N = 9$ ,  $L = 8$  e  $Q = 17$ .

### 3.2.2 Obtenção de Modelos de Aparência Utilizando Perfis RGB

Os marcadores têm que estar localizados com antecedência para treinar o ASM. Os pontos PDM são procurados em torno da vizinhança dos marcadores detectados (i.e., no

perfil do marcador), ambos, nas imagens de treinamento e nas imagens de entrada. A seguir, apresenta-se o procedimento para o cálculo dos vetores derivada do perfil normalizado aplicado às imagens a cores de um objeto para a busca de marcadores em tais imagens.

Uma estrutura do perfil é amostrada com  $\epsilon$  *pixels* por canal  $h$  de cor em torno da posição provável de cada marcador na imagem de treinamento. No  $q$ -ésimo marcador, o vetor do perfil é:

$$\mathbf{f}_{jq_h} = [f(1)_{jq_h}, \dots, f(\epsilon)_{jq_h}], \quad (19)$$

onde  $\mathbf{f}_{jq_h}$  é o vetor do perfil no  $h$ -ésimo canal de cor da imagem, e  $f(\cdot)_{jq_h}$  é o nível de cinza do *pixel* da vizinhança do  $q$ -ésimo marcador no  $h$ -ésimo canal de cor da imagem. O vetor do perfil é centrado no ponto provável de localização do marcador, e inclui os níveis de cinza dos *pixels* capturados ao longo da normal no contorno local do marcador (COOTES; TAYLOR, 1993; WAN; LAM; NG, 2005).

O vetor derivada do perfil no  $q$ -ésimo marcador no canal  $h$  de cor, é expresso pelas mudanças dos níveis de cinza que ocorrem entre *pixels* adjacentes no vetor do perfil indicado na Equação (19):

$$\mathbf{df}_{jq_h} = [f(2)_{jq_h} - f(1)_{jq_h}, \dots, f(\epsilon)_{jq_h} - f(\epsilon - 1)_{jq_h}], \quad (20)$$

onde  $\mathbf{df}_{jq_h}$  denota o vetor derivada do perfil no  $h$ -ésimo canal de cor, o qual também pode representar compactamente os canais de cores  $h = 1, \dots, \bar{h}$  (para o modelo de cor *Red Green Blue* (RGB),  $\bar{h} = 3$ ):

$$\mathbf{df}_{jq} = [\mathbf{df}_{jq_1} | \mathbf{df}_{jq_2} | \mathbf{df}_{jq_3}]^T, \quad (21)$$

o qual denota os três vetores derivada do perfil (um por cada canal), concatenados como uma matriz com uma única coluna (uma coluna concatena todos os três vetores derivada do perfil), calculado para o  $q$ -ésimo marcador da  $j$ -ésima imagem do treinamento. O vetor derivada do perfil normalizado denotado  $\mathbf{g}_{jq}$  é definido como:

$$\mathbf{g}_{jq} = \frac{\mathbf{df}_{jq}}{\sum_{b=1}^{3(\epsilon-1)} |df(b)_{jq}|}, \quad (22)$$

onde  $|df(b)_{jq}|$  é o valor absoluto do  $b$ -ésimo elemento do vetor derivada do perfil associado ao  $q$ -ésimo marcador.

A variabilidade dos vetores derivada do perfil normalizado é modelada para diferentes marcadores no conjunto de dados de treinamento  $\mathbf{X}$ , e com base nessa informação de variabilidade, o modelo do objeto é ajustado para fazer a comparação com a forma do objeto de uma imagem de entrada. As variações do perfil podem ser modeladas utilizando análise de componentes principais (PCA), como indicado na Equação (18). O modelo de aparência do objeto é treinado com  $J$  imagens de treinamento (com os marcadores colocados com antecedência), representando a forma do objeto em diferentes posições na imagem. Para o  $q$ -ésimo marcador um conjunto de  $J$  vetores derivada do perfil normalizado (um por imagem de treinamento) estão dispostas em uma matriz, em que as colunas são os  $J$  vetores derivadas do perfil normalizado:

$$\mathbf{G}_q = [\mathbf{g}_{1q}, \mathbf{g}_{2q}, \dots, \mathbf{g}_{Jq}], \quad (23)$$

onde  $\mathbf{G}_q$  é a matriz chamada simplesmente de matriz de perfis no  $q$ -ésimo marcador.

Portanto, o vetor derivada do perfil normalizado (chamado aqui simplesmente como modelo do perfil) ao redor do  $q$ -ésimo marcador pode ser representado com base em um modelo semelhante à Equação (18), como a seguir:

$$\hat{\mathbf{g}}_q = \bar{\mathbf{g}}_q + \mathbf{W}_q \mathbf{r}_q, \quad (24)$$

onde  $\hat{\mathbf{g}}_q$  é o modelo do perfil no  $q$ -ésimo marcador (i.e. o modelo do perfil obtido com o modelo e o parâmetro  $\mathbf{r}_q$ ),  $\bar{\mathbf{g}}_q$  é a média do modelo do perfil para o  $q$ -ésimo marcador de tamanho  $h(\epsilon - 1) \times J$  (construído com  $J$  imagens de treinamento como  $\bar{\mathbf{g}}_q = (1/J) \sum_{j=1}^J \mathbf{g}_{jq}$ ). A matriz  $\mathbf{W}_q$  contém  $\kappa$  autovetores da matriz de covariância de  $\mathbf{G}_q$ , associado aos autovalores  $\gamma_\iota$ , ( $\iota = 1, \dots, \kappa$ ), então  $\mathbf{W}_q \in \mathbb{R}^{h(\epsilon-1) \times \kappa}$ . O vetor de parâmetros  $\mathbf{r}_q$  de tamanho  $\kappa \times 1$ , ajusta o modelo do perfil para coincidir com o vetor do

perfil local na imagem de entrada, ao redor do  $q$ -ésimo marcador (i.e., que minimize a norma Euclidiana de uma diferença entre vetores  $\|\hat{\mathbf{g}}_q - \mathbf{g}_q\|_2^2$ ). Para reduzir a chance de detectar pontos espúrios de feições restringe-se as variações do  $\iota$ -ésimo parâmetro  $r_\iota$  em  $\pm 3\sqrt{\gamma_\iota}$ . O vetor de parâmetros do perfil na Equação (24) é dado por  $\mathbf{r}_q = \mathbf{W}_q^\top (\hat{\mathbf{g}}_q - \bar{\mathbf{g}}_q)$  (COOTES; TAYLOR, 1993).

Agora, considerando-se todos os marcadores simultaneamente, o modelo de aparência ou modelo dos perfis, que é utilizado para representar o objeto na sua excursão ao longo da imagem, pode ser descrito como:

$$\hat{\mathbf{g}} = \bar{\mathbf{g}} + \mathbf{W}\mathbf{r}, \quad (25)$$

onde  $\hat{\mathbf{g}} = [\hat{\mathbf{g}}_1^\top, \dots, \hat{\mathbf{g}}_Q^\top]^\top$  é o modelo de aparência ou modelo dos perfis que representa os perfis de todos os  $Q$  marcadores,  $\bar{\mathbf{g}} = [\bar{\mathbf{g}}_1^\top, \dots, \bar{\mathbf{g}}_Q^\top]^\top$  é a média do modelo dos perfis para todos os  $Q$  perfis nos marcadores,  $\mathbf{W} = \text{diag}[\mathbf{W}_1, \dots, \mathbf{W}_Q]$  é a matriz diagonal de autovetores  $\mathbf{W}_q$  considerando todos os  $Q$  perfis nos marcadores, e  $\mathbf{r} = [\mathbf{r}_1^\top, \dots, \mathbf{r}_Q^\top]^\top$  é o vetor de parâmetros do modelo dos perfis (que representa os perfis de todos os  $Q$  marcadores).

### 3.2.3 Localização dos Marcadores Utilizando os Perfis RGB

Nesta abordagem procura-se representar o objeto utilizando um ASM treinado para a estimativa da forma do objeto, dadas umas estimativas confiáveis dos  $Q$  marcadores.

Um processo de comparação dos perfis é utilizado para localizar todos os  $Q$  marcadores na imagem de entrada, cuja estrutura de perfis é representada compactamente na Equação (25). Para o  $q$ -ésimo marcador, o perfil estimado é  $\hat{\mathbf{g}}_q$ . A forma e os parâmetros do modelo de aparência são obtidos após a comparação dos perfis para todos os  $Q$  marcadores. Para o vetor de parâmetros de forma  $\mathbf{m}_z$ , existem  $\varsigma$  elementos por instância, (que é o número de autovetores da matriz de covariância de  $\mathbf{X}$ , veja-se a Equação (18)) e para o vetor de parâmetros de aparência  $\mathbf{r}_q$ , existem  $\kappa$  elementos, (onde  $\kappa$  é o número de autovetores da matriz de covariância de  $\mathbf{G}_q$ , veja-se a Equação (24)).

Inicialmente, os marcadores são procurados na imagem de entrada do objeto. A busca pelos marcadores é inicializada nas normais das bordas (BREJL; SONKA, 2000), e o

procedimento de comparação da forma prossegue fazendo deslocamentos a partir da estimativa da posição inicial do marcador ( i.e., a partir de  $\varrho = 0$  ), ao longo da normal da borda com  $\varrho \in [-\eta, \eta]$ . O procedimento de busca do marcador pode ser sumarizado como a seguir: dado um  $t$ -ésimo conjunto PDM de teste, (correspondente à  $t$ -ésima imagem de teste), a posição do  $q$ -ésimo marcador candidato é indicado pelo parâmetro  $\varrho$  que minimiza a função de custo:

$$\rho_q = \underset{\varrho \in [-\eta, \eta]}{\operatorname{argmin}} \|\mathbf{g}_q^{t(\varrho)} - \hat{\mathbf{g}}_q\|_2^2 \quad (26)$$

onde  $\mathbf{g}_q^{t(\varrho)}$  é o vetor derivada do perfil normalizado de uma  $t$ -ésima imagem de teste, no  $q$ -ésimo marcador deslocado  $\varrho$  posições ao longo do intervalo  $[-\eta, \eta]$ ,  $\hat{\mathbf{g}}_q$  é o modelo do perfil expresso na Equação (24),  $\|\cdot\|_2$  é a norma  $\ell_2$  ou norma Euclidiana,  $\rho_q$  é a distância a partir do  $q$ -ésimo marcador até a posição do marcador candidato. Note que a norma Euclidiana quadrática na função de custo na Equação (26) é o erro quadrado em notação de vetores, então a minimização na Equação (26) é o erro mínimo quadrado encontrado durante a comparação. Uma vez que a comparação é feita para todos os marcadores é possível atualizar os parâmetros do modelo de forma através do deslocamento. O deslocamento é a mudança de posição do marcador da posição original para um local onde se espera ter um marcador. O deslocamento é realizado após o procedimento de comparação ter sido feito para todos os marcadores. O vetor de deslocamento em uma  $z$ -ésima instância, denotado  $\Delta_{\mathbf{x}_z}$ , é dado por

$$\Delta_{\mathbf{x}_z} = [x_{\rho_1}, y_{\rho_1}, \dots, x_{\rho_Q}, y_{\rho_Q}]^T, \quad (27)$$

onde  $(x_{\rho_q}, y_{\rho_q})$  são as coordenadas da posição candidata do  $q$ -ésimo marcador. A Figura 7 ilustra o espaço para a localização do  $q$ -ésimo marcador onde uma busca local segue a direção normal.

Após o procedimento de comparação, tanto os parâmetros do modelo de forma quanto os parâmetros do modelo de aparência devem ser atualizados conforme a restrição de variabilidade (i.e.  $\pm 3\sqrt{\lambda_l}$  no caso do modelo de forma e  $\pm 3\sqrt{\gamma_l}$  no caso do modelo de

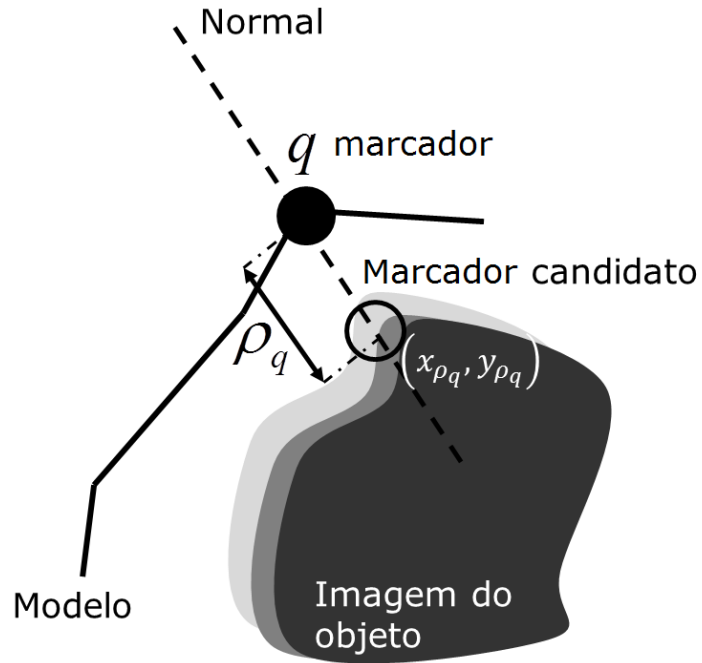


Figura 7: Ilustração de um espaço para a localização do  $q$ -ésimo marcador onde uma busca local segue a direção normal.

aparência). Considerando-se a estimativa do vetor do deslocamento na Equação (27), e o modelo de forma na Equação (18), o vetor de parâmetros de forma é atualizado conforme

$$\mathbf{m}_z = \mathbf{H}^\top (\Delta_{x_z} - \bar{\mathbf{x}}). \quad (28)$$

O vetor de parâmetros do perfil é atualizado para cada  $q$ -ésimo marcador como

$$\mathbf{r}_q = \mathbf{W}_q^\top (\mathbf{g}_q^{t(\rho_q)} - \bar{\mathbf{g}}_q), \quad (29)$$

onde  $\mathbf{g}_q^{t(\rho_q)}$  é o vetor derivada do perfil normalizado da  $t$ -ésima imagem de teste, no  $q$ -ésimo marcador deslocado  $\rho_q$  posições. O valor do  $\rho_q$  é o resultado da minimização da função custo da Equação (26) encontrada durante a comparação. O modelo de aparência na Equação (25) é preenchido com cada um dos parâmetros atualizados.

## 4 RESULTADOS EXPERIMENTAIS E DISCUSSÃO

Neste capítulo apresenta-se os resultados experimentais e discussão das inovações propostas nos modelos ativos de formas, que correspondem às contribuições na melhora do ASM para o reconhecimento de objetos e modificações do ASM na melhora da comparação de objetos em imagens a cores. Os resultados experimentais ilustrados na seção 4.1 e na seção 4.2 são dados a partir de aplicações do método proposto na melhora do ASM para o reconhecimento de objetos apresentado na seção 3.1. Os resultados experimentais ilustrados na seção 4.3 são dados a partir de uma aplicação do método proposto do ASM modificado na melhora da comparação de objetos em imagens a cores apresentado na seção 3.2.

### 4.1 Aplicações em Reconhecimento de Faces

Para a avaliação de desempenho do método proposto na seção 3.1 contextualizado para aplicações em reconhecimento de faces (i.e. com  $\Omega_{ck}$  como critério para o reconhecimento de faces), foram testados: a abordagem ASM padrão; dois métodos conhecidos adequados para redução de dimensionalidade, a saber, a análise de componentes principais (PCA) (TURK; PENTLAND, 1991), e análise de regressão discriminante espectral (SRDA) (DENG; XIAOFEI; JIAWEI, 2008); e o método de fusão melhorada de dados para reconhecimento de faces (DFBFR) (DEMIREL; ANBARJAFARI, 2009), que apresenta um desempenho imponente e representa o estado da arte em termos de reconhecimento de faces. Esses métodos foram testados na base de dados de faces Essex (ESSEX, 2011), cujas imagens contém uma diversidade grande de usuários ( foram utilizadas 100



classes que representam os indivíduos; para cada classe foram utilizadas as 5 primeiras imagens para treinamento e as 15 últimas imagens para teste ). Adota-se o detector de face utilizado em (DEMIREL; ANBARJAFARI, 2009). O algoritmo ASM para faces foi implementado com  $N = 68$  marcadores, e  $w = 11$  *pixels*. Os dois canais de crominância  $C_r$  e  $C_b$  foram utilizados como feições da face na imagem (DEMIREL; ANBARJAFARI, 2009)( $M = 2$ ).

Para fazer uma avaliação justa do método apresentado na seção 3.1, dois testes foram dirigidos: 1) avaliação do método utilizando faces detectadas com sucesso; e 2) avaliação sistêmica do método. A avaliação do método utilizando faces detectadas com sucesso é o teste do método descontando os erros introduzidos pelas faces detectadas incorretamente. A exatidão aqui é definida como a taxa de reconhecimento média para cada classe. Os resultados na Tabela 1 coluna (a) ilustram a melhora obtida pelo método proposto quando comparado com o ASM e outros métodos. Em comparação com outros métodos, o ASM tem baixo desempenho, em termos de exatidão, mas o seu desempenho pode ser melhorado substancialmente pela abordagem de seleção de feições (*ASM melhorado*).

Tabela 1: Comparação de desempenho para o conjunto de testes (maiores valores de exatidão)

Teste de Métodos	(a) Exatidão do Método	(b) Exatidão Sistêmica
PCA (TURK; PENTLAND, 1991)	0.9033	0.8720
SRDA (DENG; XIAOFEI; JIAWEI, 2008)	0.9207	0.9000
DFBFR (DEMIREL; ANBARJAFARI, 2009)	0.9373	0.9153
ASM ( $N = 68$ )	0.9040	0.9027
ASM melhorado (BEHAINE; SCHARCANSKI, 2012a) (ótimo: $\alpha = 0.2, R = 27$ )	0.9533	0.9193

Foram utilizados  $R = 27$  marcadores em nos testes (i.e.,  $R < N/2$ ), uma vez que isto reduz o número original de marcadores enquanto  $a_i$  é maximizado. A Figura 8 ilustra  $R = 27$  marcadores selecionados (marcadores em vermelho) em uma face no teste do método, onde os maiores valores de exatidão foram encontrados.

Na avaliação sistêmica, as imagens foram selecionadas aleatoriamente, e o desempenho do detector de face tende a afetar negativamente os resultados finais. A Tabela 1

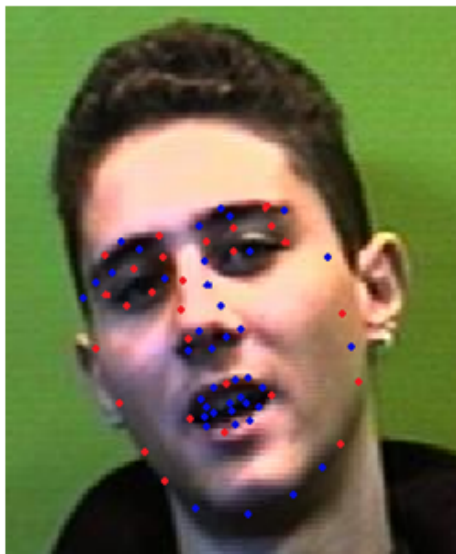


Figura 8: Ilustração de  $R = 27$  marcadores  $p_i$  (PDM) selecionados em vermelho no teste do método.

coluna (b) ilustra que o *ASM melhorado* melhora o desempenho do ASM. Isto pode ser explicado considerando que o método ASM tende a ter um desempenho pobre, quando os marcadores são afastados das bordas do quadro da imagem retornado pelo detector de face (i.e., quando as faces não são detectadas corretamente), impactando negativamente a avaliação sistêmica do ASM. O ganho do desempenho como uma função de  $\alpha$  é ilustrada na Figura 9 (a) e na Figura 9 (b), para os dois tipos de testes. Como o DFBFR não permite redução de dimensionalidade, somente é ilustrado na Tabela 1.

Outra proposta com base em algoritmos genéticos (GA), pode melhorar de forma meta-heurística os resultados do desempenho do classificador utilizado em (BEHAINE; SCHARCANSKI, 2012a), uma vez que existe uma quantidade grande de combinações para a seleção dos marcadores que otimiza a taxa de acerto (VIGNOLO et al., 2012). Em (VIGNOLO et al., 2012) um número de combinações de marcadores foi introduzido nos três processos fundamentais da GA: seleção, variação e substituição, até atingir uma exatidão de 0.98 utilizando 56 feições (localizadas em 28 marcadores utilizando os dois canais de crominância  $C_b$  e  $C_r$ ).

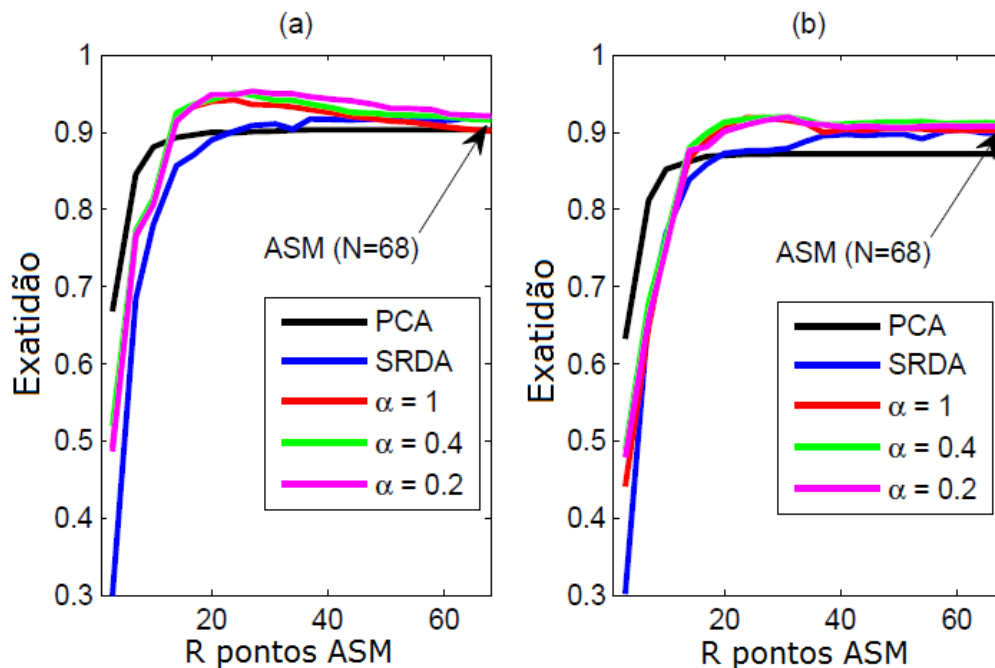


Figura 9: Comparação de desempenho utilizando: (a) Avaliação do método; e (b) Avaliação sistêmica.

## 4.2 Aplicações em Monitoramento Visual de Painéis

Nesta seção concentramo-nos na aplicação do método proposto na seção 3.1 para a seleção marcadores em ASM para o caso de leitura de painéis.

A aplicação de ASM em leitura de painéis é contextualizado no cenário onde um método automático de monitoramento visual é utilizado como ferramenta redundante para acessar remotamente medidores analógicos, e o medidor momentaneamente (ou permanentemente) não tem uma interface de comunicação digital disponível. A Figura 10 (a) ilustra as etapas de um esquema de monitoramento visual automático de painéis analógicos. O sistema redundante ilustrado na Figura 10 (a) é designado inicializar a operação quando a topologia tradicional para monitoramento de medidores falha, como ilustrado na Figura 10 (b).

Também, tem-se interesse no uso de tecnologias de painéis com baixo consumo de potência, i.e. medidores de painel analógico. A novidade nesta seção é a utilização de modelos ativos de formas (ASM) como um modelo flexível para representar as deforma-

ções de um painel analógico através de marcadores ao longo de sua escala de medição. Estes marcadores mudam a sua posição com as deformações do objeto. Utiliza-se estes marcadores para identificar automaticamente a região de interesse (ROI) em imagens de painéis analógicos, e depois é feita a leitura com base nas formas identificadas do medidor.

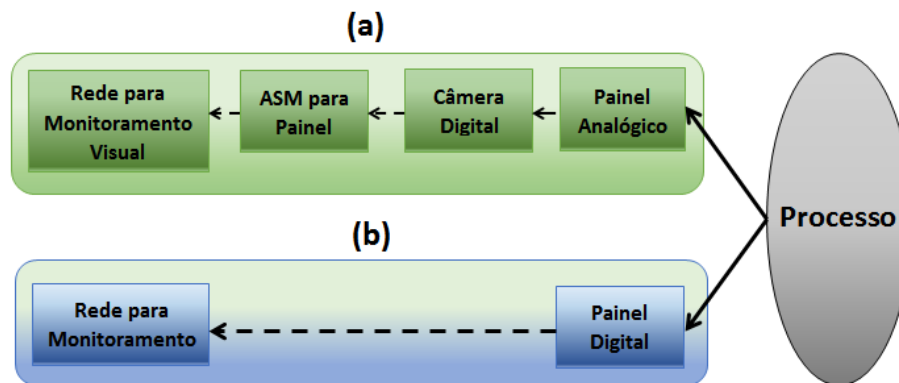


Figura 10: (a) Esquema de redundância ativa para monitoramento visual de medidores com painel analógico, e (b) Rede utilizada para o monitoramento de medidores.

Um conjunto de marcadores é definido para a captura de informação geométrica de um medidor com painel analógico dado. Define-se  $N = 14$  marcadores no PDM, como ilustrado na Figura 11 (a). O número de marcadores depende da forma do painel, e em alguns casos mais do que 14 poderiam ser utilizados. Cada marcador  $p_i$  no PDM representa uma localização relevante da forma (i.e. 11 marcadores para a caixa de bordas, e 3 marcadores para o ponteiro). O aspecto visual em cada localização do marcador na imagem do painel é descrita pelo conjunto de feições (i.e. posição  $x$ , posição  $y$ , luminância, textura, etc.). Neste caso, o ASM utiliza as posições  $(x, y)$  dos marcadores como feições, e os perfis de luminância dos marcadores como aparência.

Os painéis analógicos usualmente tem divisão por intervalos. Utiliza-se um ASM de  $N$  pontos PDM  $S_{k,\epsilon}$  para representar os valores de leitura na  $k$ -ésima forma do painel. Associa-se cada  $S_{k,\epsilon}$  à leitura ao longo da escala do painel analógico,  $k = 1, \dots, K$ , denotado como  $S_{k,\epsilon} = \{p_i(x_i + \epsilon_{x_i}, y_i + \epsilon_{y_i})\}^k$ ,  $i = 1, \dots, N$ , onde  $(x_i, y_i)$  são as coordenadas de um ponto do PDM  $p_i$  e  $(\epsilon_{x_i}, \epsilon_{y_i})$  são os erros de localização. Para limitar a variabilidade da forma, divide-se o painel do medidor em  $K$  intervalos (i.e. para  $K = 3$  tem-se

3 ASMs, um  $k$ -ésimo ASM para cada intervalo do painel do medidor). Cada ASM é treinado com  $J$  imagens de treinamento que pertencem ao seu intervalo de medição. Os valores das medidas obtidas pelo dispositivo de referência são associadas às formas dos PDMs (inclinação do ponteiro) em cada uma das  $J$  imagens de treinamento por classe. A seleção dos marcadores do PDM para a leitura automática de medidores com painel analógico é detalhada a seguir.

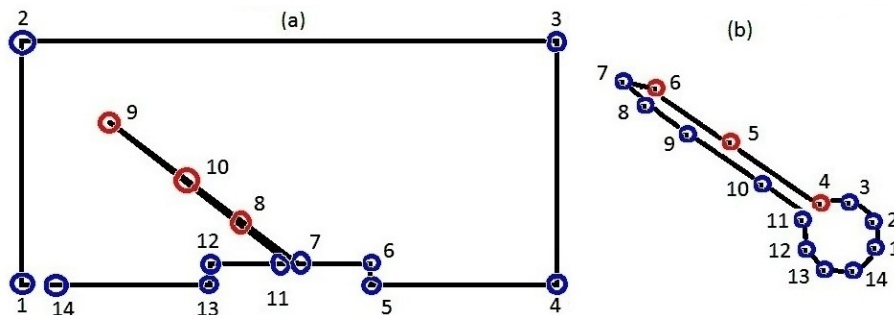


Figura 11: (a) Ilustração de marcadores  $p_i$  (PDM) utilizados no ASM proposto para o painel analógico, com  $N = 14$  e  $R = 3$  marcadores principais localizados para a ROI, e (b) Outra forma do PDM em um ASM para um painel analógico.

Em um ASM para painel analógico tem-se  $(N - R)$  marcadores que não são úteis para a leitura, (i.e. a caixa do painel) e outros que são úteis para a leitura, (e.g. o ponteiro do painel). Foca-se na região de interesse (ROI), que tem marcadores úteis para a leitura do painel (com o PDM do ASM). Os marcadores relevantes foram calculados segundo o procedimento ilustrado na seção 3.1 com o PDM ilustrado na Figura 12 (b), onde a função de pesos  $a_i$  é calculada a partir da Equação (14) para as  $K = 3$  escalas do painel representada por cada  $S_k$  com  $k = 1, \dots, K$  e  $J = 5$  imagens de treinamento. As feições relevantes foram identificadas como as posições do ponteiro, (posição no eixo  $x$  e no eixo  $y$ , i.e.  $(x_i, y_i)$  de um marcador  $p_i$ , então  $M = 2$ ). Por exemplo, para um número reduzido de marcadores  $R_e = 5$ , a função de pesos identifica os marcadores relevantes do painel como é ilustrado na Figura 12 (a) (valores em vermelho). Os valores de  $a_i$  na Figura 12 (a) sugerem que no mínimo são necessários  $R = 3$  marcadores localizados no ponteiro para realizar um reconhecimento de classes.

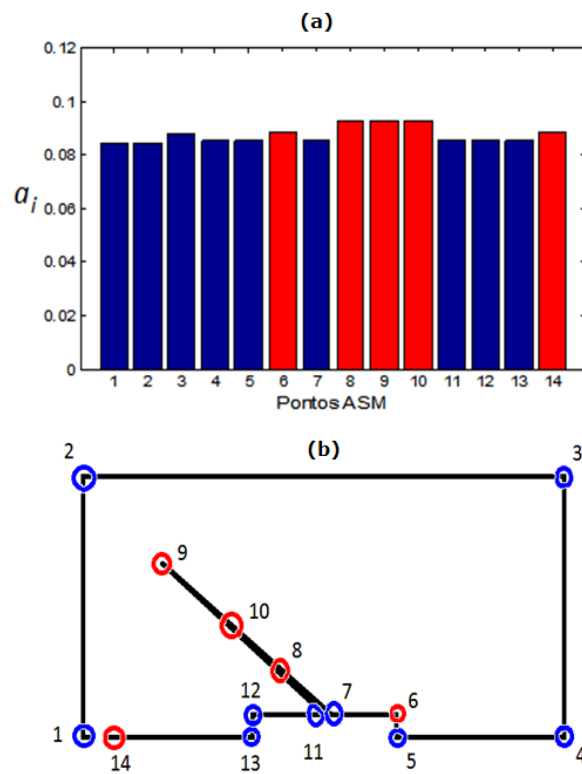


Figura 12: (a) Exemplo de valores da função de pesos  $a_i$  com  $R_e = 5$  no PDM de um painel analógico (marcadores selecionados em vermelho para um valor  $R_e = 5$ ), e a sua (b) Ilustração no PDM respectivo (marcadores selecionados em círculos vermelhos vazios para um valor  $R_e = 5$ ).

Utiliza-se  $R = 3$  em nos testes porque é o número de marcadores no ponteiro do painel como ilustrado na Figura 11 (a). A  $k'$ -ésima classe dos ASMs é reconhecida pela expressão:

$$k' = \arg \min_{\forall k \in \{K\}} \{|\hat{I}^R - \hat{S}_k^R|^2\}, \quad (30)$$

onde  $\hat{I}^R$  representa o valor esperado do vetor perfil (luminância) nos  $R$  marcadores selecionados durante o treinamento, e  $\hat{S}_k^R$  são os valores de luminância localizados nos  $R$  marcadores do  $k$  étimo ASM selecionados durante a comparação, com  $k = 1, \dots, K$ . Nos testes  $K = 3$  representa o número de intervalos de medição,  $k = 1$  para intervalo baixo,  $k = 2$  para intervalo médio e  $k = 3$  para intervalo alto. Desde que o intervalo do medidor seja reconhecido, a inclinação do ponteiro representado pelos  $R$  marcadores, poderia ser calculada utilizando uma regressão linear durante o treinamento. A leitura do painel analógico é obtida por:

$$\vartheta = \beta \dot{m}(S_{k'}^R) + \lambda, \quad (31)$$

onde  $\vartheta$  é o valor da leitura,  $\dot{m}(S_{k'}^R)$  é a inclinação do ponteiro indicada pelos  $R$  marcadores selecionados do ponteiro localizado na  $k'$ -ésima classe ASM, e  $\lambda, \beta$  são os parâmetros da regressão linear, calculadas pelo estimador de menor diferença quadrática (REIS, 2008).

O parâmetro fundamental em um painel analógico é a escala mínima de divisão  $\check{\epsilon}$ , que limita a precisão (tolerância) em um monitoramento visual de um painel analógico em  $\delta = \check{\epsilon}/2$ . Nos testes foi utilizado o medidor HIOKI MODEL 3007, que tem um painel analógico com intervalo  $[0, 12]$  volts DC e uma escala mínima de divisão  $\check{\epsilon} = 0.2$  volts, como ilustrado na Figura 13 (a). O intervalo  $[0, 12]$  volts DC da imagem do painel analógico foi representado com 184 *pixels*, com uma resolução aproximada de 3 *pixels* na escala mínima de divisão  $\check{\epsilon} = 0.2$  volts DC. O sensor da câmera digital tem uma resolução de  $(640 \times 480)$  *pixels* por imagem. O ASM para o painel analógico foi implementado para  $N = 14$  marcadores, e  $R = 3$  pontos ASM foram selecionados para a ROI, como é ilustrado na Figura 11 (a). O ASM foi treinado utilizando 5 imagens do painel para

cada  $k$ -ésima classe do ASM. A classe do modelo associado com  $K = 3$  foi,  $k = 1$  no intervalo  $[0, 4)$  volts DC,  $k = 2$  no intervalo  $[4, 8)$  volts DC, e  $k = 3$  no intervalo  $[8, 12]$  volts DC. O algoritmo ASM foi implementado seguindo o procedimento em (COOTES, 2000). O esquema ASM proposto para o monitoramento de um painel analógico foi implementado em *Matlab*. A média do tempo de execução do esquema ASM proposto foi 24.66 segundos, utilizando um computador IBM-PC compatível com um processador Intel de dois núcleos de 2.67 Mhz e 8 Gb de RAM. Os resultados da comparação obtidos com o medidor com painel modelado com o nosso ASM é ilustrado na Figura 13 (b).

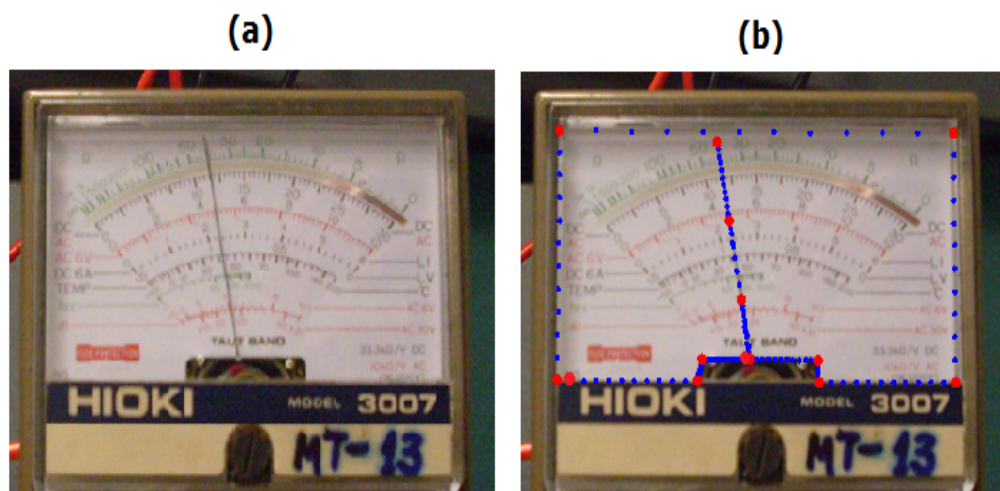


Figura 13: (a) Painel analógico do medidor HIOKI MODEL 3007, e (b) Comparação do ASM para o painel analógico do medidor HIOKI MODEL 3007, marcadores em pontos vermelhos e bordas em pontos azuis.

O método proposto apresentado em (BEHAINE; SCHARCANSKI, 2012b) para o monitoramento visual de painéis em medidores, foi testado com 15 imagens do painel analógico ( $640 \times 480$  pixels) para cada um dos intervalos  $[0, 12]$  volts DC e  $[0, 200]$  volts DC. A Figura 14 ilustra o desempenho da comparação do ASM, em termos da média do erro de localização dos marcadores ( o erro de localização dos marcadores é a diferença absoluta entre a posição estimada dos marcadores no modelo e a posição real dos marcadores do objeto na imagem), para o painel analógico do medidor, baseado em  $R$  marcadores  $\{p_8, p_9, p_{10}\}$  para cada classe  $\{k = 1, k = 2, k = 3\}$  do ASM no intervalo



[0, 12] volts DC. A média do erro absoluto (ou simplesmente erro) da comparação ASM é  $\leq 3$  pixels para a maioria dos  $R$  marcadores, que é igual (ou menor do que) o número de pixels da escala mínima de divisão  $\check{e}$  (3 pixels em  $\check{e} = 0.2$  volts DC), fornecendo a resolução de leitura suficiente considerando que esteja livre de erros geométricos ou paralaxe. As medições do método proposto no intervalo [0, 12] volts DC com relação à referência são ilustradas na Figura 15.

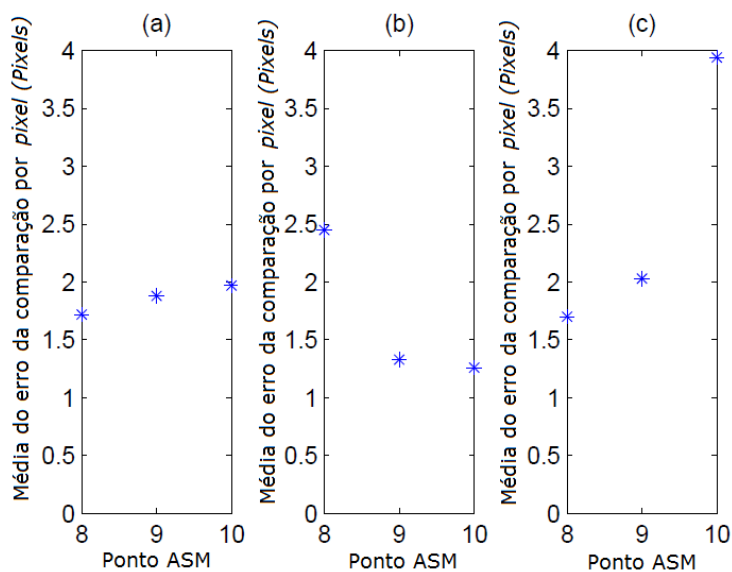


Figura 14: Média do erro da comparação ASM por *pixel* obtido no painel analógico, (a) Modelo classe  $k = 1$ , (b) Modelo classe  $k = 2$ , e (c) Modelo classe  $k = 3$ .

A exatidão é o erro da medida quando comparado com a medida fornecida a partir um dispositivo utilizado na calibração (dispositivo referência) (KAUFMAN; SEIDMAN, 1988). O dispositivo referência utilizado foi o multímetro digital Minipa modelo ET 2615A (LCD de 4 dígitos). A exatidão foi representada pelo desvio padrão do erro absoluto  $\sigma$  e a média do erro absoluto  $\mu$ . A Tabela 2 parte (a) ilustra a exatidão da medição para o esquema proposto de monitoramento visual de medidores considerando-se a tolerância da medição visual  $\mu \leq \delta$ . O erro absoluto mínimo é zero, e os erros absolutos máximos estão ilustrados também na Tabela 2 parte (a). Os dados visuais foram captura-

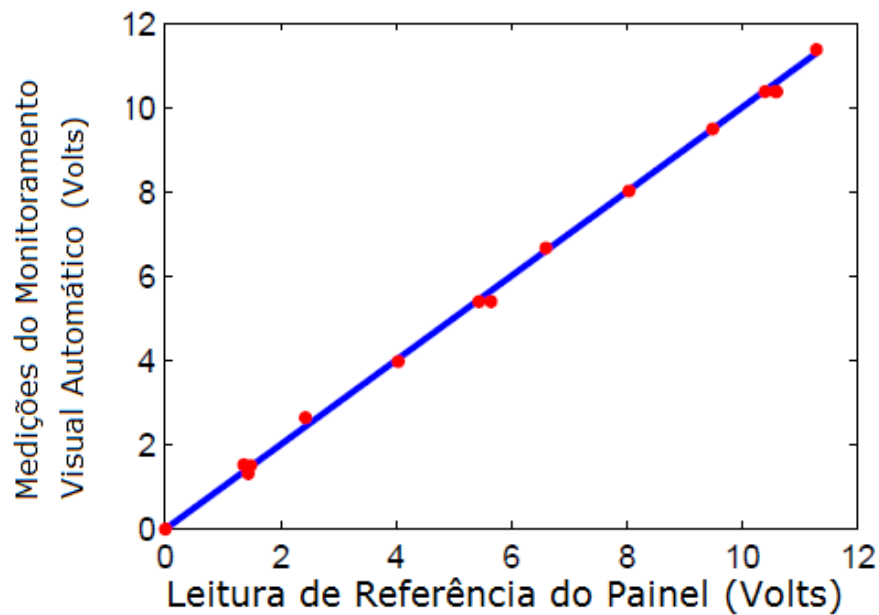


Figura 15: Medições do método do monitoramento automático visual apresentado (Volts).

dos em diferentes distâncias do observador ( $d = 60$  cm e  $d = 30$  cm, com duas pessoas) para ilustrar o efeito da resolução nos testes. A Tabela 2 parte (a) ilustra a estatística da exatidão para o esquema do ASM apresentado quando comparado com a exatidão da leitura humana. O desempenho do esquema ASM apresentado em termos do desvio padrão (0.0891 volts DC) é parecido ao desempenho de uma leitura humana no medidor desde uma distância de  $d = 60$  cm (0.0803 volts DC), mas não tem um desempenho tão bom quando comparado com uma leitura humana capturada desde uma distância de  $d = 30$  cm (0.0723 volts DC), isto é possivelmente, porque em distância curta ( $d = 30$  cm), a resolução da imagem é menor do que a imagem vista pela leitura humana (BEHAINE; SCHARCANSKI, 2012b). A imagem do painel analógico é também degradada pelo formato para compressão com perda *Joint Photographic Experts Group* (JPEG). Entretanto, o desempenho do método proposto em termos da média (0.1020 volts DC) foi superior quando comparado com a leitura humana em todas as imagens testadas. Os resultados ilustrados na Tabela 2 parte (b) sugerem que a condição de tolerância  $\mu \leq \delta$  (1.2327 volts

DC  $\leq 2.5$  volts DC) é satisfeita quando é utilizada uma resolução alta (i.e. mais *pixels* utilizados para representar cada divisão de escala  $\check{e}$ ).

Tabela 2: Estatística da exatidão para o esquema apresentado de monitoramento visual de medidores com painel, (a) 3 *pixels* por  $\check{e}$  e (b) 9 *pixels* por  $\check{e}$ .

(a) Intervalo $[0, 12]$ Volts DC, $\delta = 0.1$ Volts DC	Erro Absoluto (Volts DC)			
	$\mu$	$\sigma$	Min	Max
ASM (BEHAINE; SCHARCANSKI, 2012b) $d = 30$ cm	0.1020	0.0891	0.00	0.26
Humano $d = 30$ cm	0.1460	0.0723	0.00	0.24
Humano $d = 60$ cm	0.1060	0.0803	0.00	0.34
(b) Intervalo $[0, 200]$ Volts DC, $\delta = 2.5$ Volts DC	$\mu$	$\sigma$	Min	Max
ASM (BEHAINE; SCHARCANSKI, 2012b) $d = 30$ cm	1.2327	0.9380	0.03	2.93
Humano $d = 30$ cm	1.9000	1.1958	0.20	3.80
Humano $d = 60$ cm	1.9933	1.2062	0.20	4.20

Como mencionado anteriormente, o esquema ASM apresentado na seção 4.2 pode ser adaptado para diferentes formas de painéis analógicos. Com o intuito de ilustrar o comportamento do ASM proposto em outros tipos de painéis analógicos, foi testado o PDM ilustrado na Figura 11 (b), cujos resultados da comparação estão ilustrados na Figura 16. A Tabela 3 ilustra a estatística da exatidão para o esquema apresentado de monitoramento visual do painel ilustrado na Figura 16, cujos resultados satisfazem a condição de tolerância  $\mu \leq \delta$  (1.4109 volts AC  $\leq 2.5$  volts AC). Se uma imagem ilustrar diferentes medidores com painéis, dois ou mais ASMs poderiam ser utilizados (um para cada tipo de painel que se pretende monitorar), como vai ser apresentado com detalhes na seção 4.3, mas utilizando um ASM modificado proposto na seção 3.2.

Tabela 3: Estatística da exatidão para o esquema apresentado de monitoramento visual do painel ilustrado na Figura 16, 5 *pixels* por  $\check{e}$ .

Intervalo $[50, 200]$ Volts AC, $\delta = 2.5$ Volts AC	Erro Absoluto (Volts AC)			
	$\mu$	$\sigma$	Min	Max
ASM (BEHAINE; SCHARCANSKI, 2012b) $d = 30$ cm	1.4109	1.2248	0.19	4.02
Humano $d = 30$ cm	1.7333	1.7915	0.00	7.00
Humano $d = 60$ cm	2.5333	1.9591	0.00	8.00

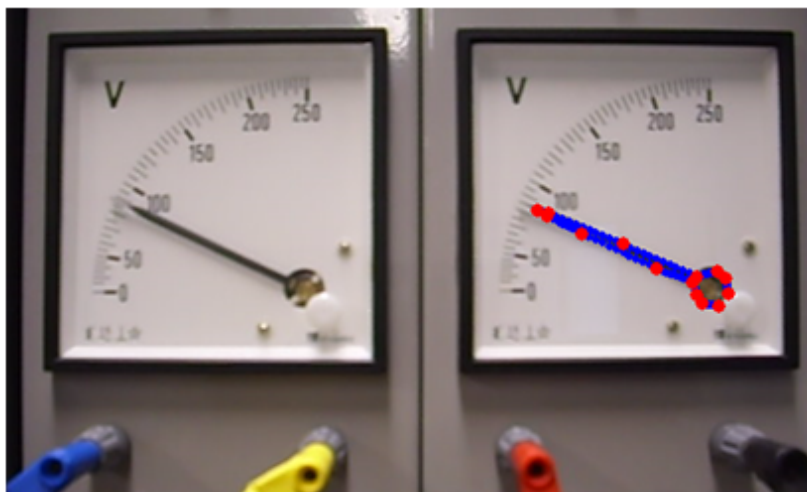


Figura 16: Comparação do ASM em um painel analógico com um PDM ilustrado na Figura 11 (b), marcadores em pontos vermelhos e bordas em pontos azuis.

### 4.3 Aplicações em Monitoramento Visual de Multipainéis

Nesta seção concentramo-nos na aplicação do método proposto na seção 3.2 para a comparação dos marcadores adequados em um PDM (já com seleção de marcadores seguindo o método proposto na seção 3.1, contextualizado na seção 4.2) que representam os ponteiros que permitem a leitura visual de multipainéis.

A leitura visual automática de um medidor com painel analógico individual foi introduzida em (BEHAINE; SCHARCANSKI, 2012b). Tem-se duas fases neste processo, o treinamento ASM e a comparação do painel (fase de leitura). Os dois estágios, o treinamento e a comparação, são realizados para cada subclasse ASM correspondente a um subintervalo da escala completa do painel analógico. Durante o treinamento dois modelos são construídos. Um modelo que captura a forma de um conjunto de coordenadas de marcadores, tal como explicado na seção 3.2.1 e indicado na Equação (18). O segundo modelo é construído a partir dos perfis como explicado na Seção 3.2.2 e Equação (25). Os valores de leitura de referência são fornecidos pela calibração e as imagens de treinamento são marcadas com as leituras obtidas durante o treinamento. Durante a comparação da forma em uma imagem de entrada, a leitura visual automática do ponteiro é obtida di-

retamente a partir da inclinação dos marcadores associados ao ponteiro do medidor. Nos materiais utilizados na seção 4.2 apenas os "marcadores identificados"  $\{P_{a_n}\}$  são usados para a estimativa da inclinação do ponteiro utilizando um modelo de regressão linear (REIS, 2008).

Na seção 3.2 usa-se um esquema PDM especial, o que implica em um esquema diferente para a definição de marcador, representação do perfil e localização do marcador, como explicado nas seções 3.2.1, 3.2.2 e 3.2.3, respectivamente. Utiliza-se aqui "marcadores identificados" e "marcadores interpolados" como ilustrado na Figura 17 (a) e na Figura 17 (b). Os "marcadores identificados" estão localizados manualmente para cada imagem do painel exibida durante a fase de treinamento, e os "marcadores interpolados" estão localizados automaticamente por interpolação linear entre os "marcadores identificados".

O procedimento da comparação no ASM visa encontrar a melhor posição e os melhores parâmetros da forma que se encaixam em uma instância do modelo  $\mathbf{x}_z$ , inicializado com o valor esperado, para um novo conjunto de pontos da imagem  $\Delta_{\mathbf{x}_z}$ . O espaço da comparação é avaliado na representação do modelo de aparência, calculado de acordo com o procedimento na seção 3.2.3. Uma vez que o melhor ajuste é calculado em termos do erro mínimo quadrado da comparação, o deslocamento dos marcadores sobre a imagem é calculado pela Equação (27). O modelo de forma e o modelo de aparência são atualizados utilizando a Equação (28) e a Equação (29), respectivamente. Nos testes foi utilizado o máximo número de autovetores para cada um dos modelos. Este objetivo é atingido de forma progressiva até um número de iterações, minimizando a projeção entre o modelo correspondente e a localização dos perfis dos marcadores na imagem  $|\Delta_{\mathbf{x}_z} - T_{\tau,\xi,\theta}(\bar{\mathbf{x}} + \mathbf{H}\mathbf{m}_z)|^2$ , onde  $T_{\tau,\xi,\theta}(\bar{\mathbf{x}} + \mathbf{H}\mathbf{m}_z)$  denota as posições do PDM na representação da imagem  $\mathbf{x}_z$ , realizando uma rotação por  $\theta$ , um dimensionamento por  $\xi$ , e um deslocamento de coordenadas por  $(x, y)_\tau$ .

O esquema ASM proposto na seção 3.2 foi implementado em blocos concorrentes para a leitura de multipainéis, como é ilustrado na Figura 18. Cada ASM para painel é independente, porém é possível conseguir um monitoramento visual em paralelo de  $U$  painéis

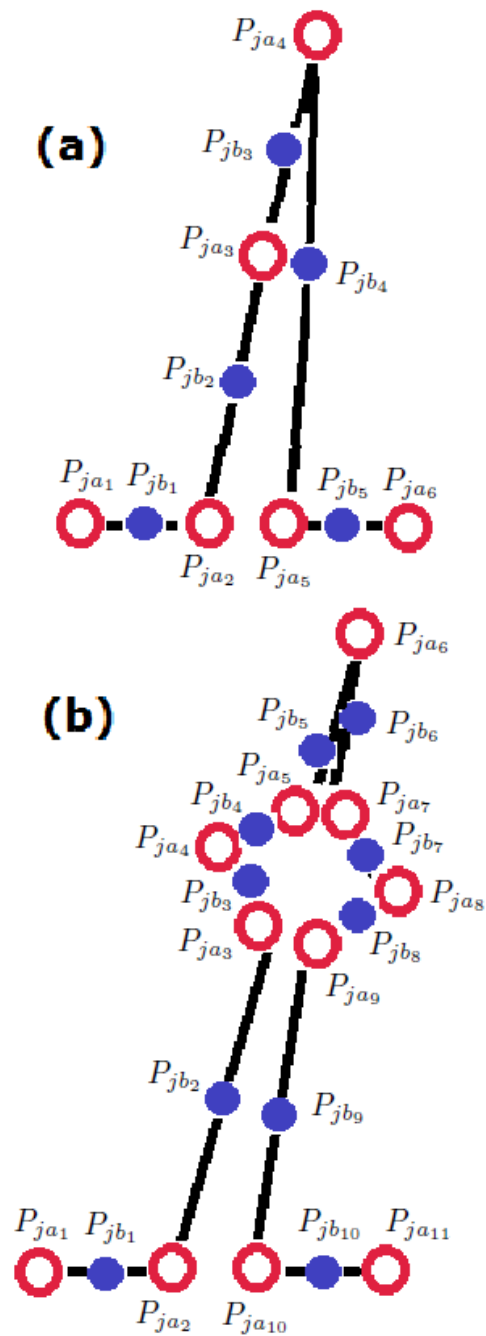


Figura 17: (a) Ilustração de um PDM utilizando marcadores identificados, (círculos vazios em vermelho) e marcadores interpolados (círculos em azul) para modelar um ponteiro analógico ASM, neste caso,  $N = 6$ ,  $L = 5$  e  $Q = 11$ , (b) Outro caso de PDM com  $N = 11$ ,  $L = 10$  e  $Q = 21$ .

analógicos. Nos experimentos foram utilizados três medidores com painéis analógicos ( $U = 3$ ). O primeiro medidor tem um painel analógico na faixa de  $[10, 40]$  volts AC, modelo RENZ BRASIL, e uma escala mínima de divisão  $\check{e} = 2$  volts AC, tal como ilustrado na Figura 19 (a) etiqueta (1). O ponteiro do painel analógico com rótulo (1) na Figura 19 (a), foi modelado pelo tipo PDM ilustrado na Figura 17 (b), e a sua leitura foi feita escolhendo os "marcadores identificados"  $\{P_{a_2}, P_{a_3}, P_{a_5}\}$ . O segundo medidor tem um painel analógico na faixa de  $[0, 50]$  volts AC, modelo MF133, e uma escala mínima de divisão  $\check{e} = 5$  volts AC, tal como ilustrado na Figura 19 (a) rótulo (2). O ponteiro do painel analógico com rótulo (2) na Figura 19 (a), foi modelado pelo tipo PDM ilustrado na Figura 17 (a), e a sua leitura foi feita escolhendo os "marcadores identificados"  $\{P_{a_2}, P_{a_3}, P_{a_4}\}$ . O terceiro medidor tem um painel analógico na faixa de  $[2, 10]$  mili amperes AC, modelo RENZ BRASIL, e uma escala mínima de divisão  $\check{e} = 0.5$  mili amperes AC, como é ilustrado na Figura 19 (a) rótulo (3). O ponteiro do painel analógico com rótulo (3) na Figura 19 (a), foi modelado pelo tipo PDM ilustrado na Figura 17 (b), e a sua leitura foi feita escolhendo os "marcadores identificados"  $\{P_{a_2}, P_{a_3}, P_{a_5}\}$ .

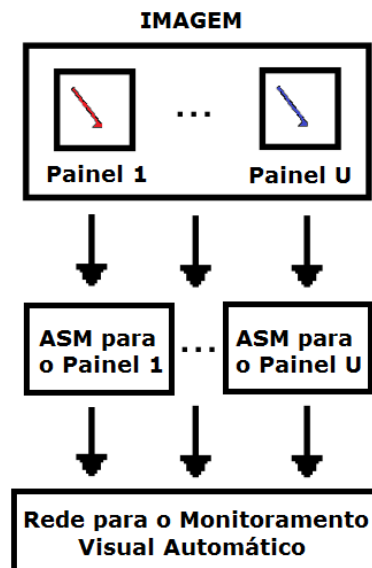


Figura 18: Ilustração de um diagrama de blocos de uma rede para monitoramento visual automático utilizando concorrentemente ASM para  $U$  painéis analógicos.

Todas as imagens dos painéis analógicos utilizam cinco subclasses ASMs uniforme-

mente espaçadas por intervalo ao longo da escala do medidor (i.e.  $S_k$  para  $k = 1, \dots, K$ , como explicado na seção 4.2, com  $K = 5$ ). Cada subclasse ASM foi treinada com 10 imagens. Todas as subclasses ASMs foram inicializadas perto da média do modelo de forma. A distância da câmera até os painéis é fixa, então  $\xi = 1$ , a rotação foi avaliada para  $\theta \in [-\pi/20, \pi/20]$ . A inicialização do ASM é uma tarefa importante quando a imagem sofre vibrações, porém uma área da imagem é verificada para cada subclasse ASM, então as coordenadas de deslocamento  $(x, y)_\tau$  são avaliadas em tal área. O tamanho de dita área é experimental, e para as condições iniciais dos painéis ilustrados na Figura 19 (a) rótulos (1), (3) e a Figura 19 (a) rótulo (2), são  $8 \times 16$ ,  $8 \times 16$  e  $4 \times 40$  *pixels* respectivamente. Os melhores resultados da comparação de cada subclasse ASM, em termos do erro mínimo quadrático, são selecionados a partir de tais áreas, para cada faixa de cada painel analógico a fim de fazer uma avaliação completa da faixa de medição.

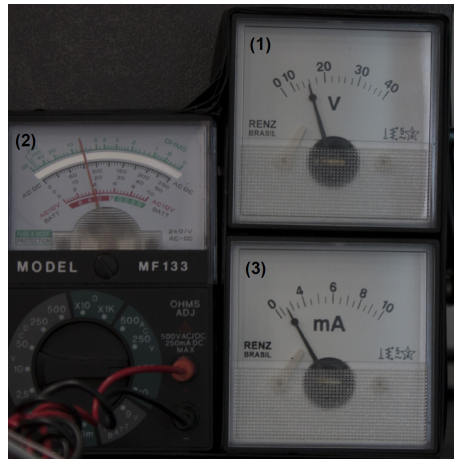
A linguagem *Matlab* foi utilizada para a implementação do esquema ASM proposto. Um computador IBM - PC compatível, com processador Intel quad core de 3.1 GHz e 32 GB de RAM foi utilizado para teste. O tempo de execução médio do esquema ASM proposto para todos os painéis testados foi 11.69 segundos. Para manter a confiabilidade e a qualidade dos testes o formato da imagem utilizado foi o *Windows Bitmap* (BMP), em vez do formato para compressão com perda *Joint Photographic Experts Group* (JPEG).

O esquema ASM proposto na seção 3.2 em aplicações de monitoramento visual automático de vários medidores com painéis analógicos foi testado com 30 imagens de painéis analógicos ( $1930 \times 1982$  *pixels*), 10 imagens por cada painel analógico em todas as faixas de medição. O sensor utilizado para adquirir as imagens foi a câmera digital de 12.2 *megapixels* (Canon EOS Rebel XSI). Os resultados da comparação do ASM proposto com  $\epsilon = 6$  *pixels*, depois de 10 iterações são ilustrados na Figura 19 (b), as leituras para os três painéis com rótulos (1), (2) e (3) são realizadas simultaneamente. Além disso, a distribuição dos diferentes tipos de marcadores estão ilustrados na Figura 19 (b), onde é fácil observar que os "marcadores identificados" são pontos relevantes do ponteiro.

A exatidão é o erro de medida quando comparado com a medida fornecida a partir um dispositivo utilizado na calibração (dispositivo referência) (KAUFMAN; SEIDMAN,



(a)



(b)

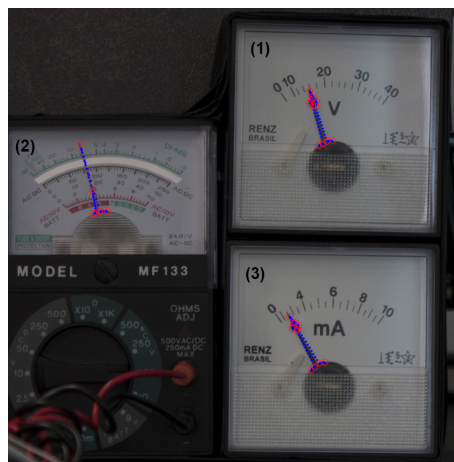


Figura 19: (a) Ilustração de uma imagem de multipainéis analógicos: (1) modelo RENZ BRASIL (AC V), (2) modelo MF133 (AC V) e (3) modelo RENZ BRASIL (AC mA); e (b) Resultados da comparação ASM com  $\epsilon = 6$  pixels, depois de 10 iterações respectivamente.

1988). O dispositivo referência utilizado foi o multímetro digital Minipa modelo ET 2615A (LCD de 4 dígitos). A exatidão foi representada pelo desvio padrão do erro absoluto  $\sigma$  e a média do erro absoluto  $\mu$ . A Tabela 4, a Tabela 5 e a Tabela 6 ilustram os resultados das estatísticas das medições utilizando os painéis analógicos com rótulos (1), (2) e (3) respectivamente, como estão ilustrados na Figura 19 (a). O parâmetro essencial em um painel analógico é a escala mínima de divisão  $\check{\epsilon}$ , o que restringe a precisão (tolerância) do monitoramento visual de painéis analógicos para  $\delta = \check{\epsilon}/2$ . No esquema de monitoramento visual automático proposto para multipainéis analógicos, foi considerado a condição de tolerância da medição visual  $\mu \leq \delta$ . O erro mínimo absoluto e o erro máximo absoluto também estão ilustrados na Tabela 4, na Tabela 5 e na Tabela 6.

A condição de tolerância  $\mu \leq \delta$  foi atingida no esquema ASM proposto para todos os tipos de painéis analógicos. Os valores das médias  $\mu$  das medidas visuais, foi mínima em todos os tipos de teste quando comparado com o método humano. Também, um teste comparativo de métodos com ASM estão ilustrados na Tabela 4, na Tabela 5 e na Tabela 6. Foram testados quatro métodos ASM. O primeiro método testado é o esquema ASM proposto. O segundo método testado para a comparação foi o esquema original ASM (COOTES et al., 1995). O método ASM original (COOTES et al., 1995), mantém o número de marcadores definidos manualmente. Os resultados do método original ASM (COOTES et al., 1995), ilustra o pior desempenho em termos do desvio padrão do erro absoluto  $\sigma$  e da média do erro absoluto  $\mu$ . O terceiro método ASM testado foi o esquema ASM não linear (VAN GINNEKEN et al., 2002). O quarto método testado foi o esquema ASM para painel individual introduzido em (BEHAINE; SCHARCANSKI, 2012b).

O posto é o processo de posicionamento de itens de estatísticas individuais ou de grupos, na escala ordinal de números em relação a outros. A medição do desempenho dos métodos ASM testados foi definida em termos do posto da média  $\mu$  nos diferentes painéis. O desempenho comparativo para cada um dos métodos testados é ilustrado na Tabela 7. Foi avaliado o desempenho do esquema ASM proposto como a média dos postos em todos os painéis analógicos  $\Omega$ , e os resultados são apresentados na Tabela 7. Em todos os testes de comparação a condição de tolerância  $\mu \leq \delta$  foi satisfeita. Neste esquema ASM

Tabela 4: Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (1).

Painel 1, $\delta = 1.0$ Volt AC Métodos ASM	Erro Absoluto (Volts AC)			
	$\mu$	$\sigma$	Min	Max
ASM proposto ( $N = 11, L = 100$ )	0.68	0.34	0.18	1.46
Humano	1.34	1.53	0.09	4.71
ASM ( $N = 11, L = 0$ )(COOTES et al., 1995)	11.95	9.92	0.47	27.46
ASM ( $N = 111, L = 0$ )(VAN GINNEKEN et al., 2002)	1.03	0.71	0.10	2.18
ASM ( $N = 111, L = 0$ )(BEHAINE; SCHARCANSKI, 2012b)	0.59	0.25	0.27	1.04

Tabela 5: Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (2).

Painel 2, $\delta = 2.5$ Volts AC Métodos ASM	Erro Absoluto (Volts AC)			
	$\mu$	$\sigma$	Min	Max
ASM proposto ( $N = 6, L = 100$ )	2.09	1.53	0.14	4.89
Humano	3.57	0.74	2.72	4.99
ASM ( $N = 6, L = 0$ )(COOTES et al., 1995)	14.29	7.78	2.12	26.16
ASM ( $N = 106, L = 0$ )(VAN GINNEKEN et al., 2002)	1.51	1.03	0.28	3.10
ASM ( $N = 106, L = 0$ )(BEHAINE; SCHARCANSKI, 2012b)	2.86	4.90	0.36	16.63

Tabela 6: Comparação da estatística da exatidão obtida com o esquema proposto para o monitoramento visual automático de multipainéis analógicos em medidores, no caso do painel com rótulo (3).

Painel 3, $\delta = 0.25$ mA AC Métodos ASM	Erro Absoluto (mA AC)			
	$\mu$	$\sigma$	Min	Max
Proposed ASM ( $N = 11, L = 100$ )	0.13	0.07	0.04	0.29
Humano	0.33	0.20	0.02	0.59
ASM ( $N = 11, L = 0$ )(COOTES et al., 1995)	3.80	3.70	0.10	14.10
ASM ( $N = 111, L = 0$ )(VAN GINNEKEN et al., 2002)	0.22	0.19	0.06	0.66
ASM ( $N = 111, L = 0$ )(BEHAINE; SCHARCANSKI, 2012b)	0.14	0.10	0.06	0.36

proposto se alcançou o melhor desempenho comparativo, em termos da média dos postos para os diferentes painéis. Por exemplo, a quantidade igual a 1, na coluna 3 (o painel 3 da Tabela 6) na Tabela 7, significa que o esquema ASM proposto atingiu o melhor  $\mu$  dentre todos os diferentes esquemas de ASM testados. Os resultados ilustrados na Tabela 7 sugerem que os métodos com base no esquema de perfil ((BEHAINE; SCHARCANSKI, 2012b) e o proposto) alcançaram as primeiras posições no posto em painéis com ponteiros em preto e fundo em branco, provavelmente porque neste caso, o contraste das bordas dos ponteiros com o fundo em tais painéis analógicos estão bem definidas, como ilustrado na Figura 19 (a) painel com rótulo (1) e painel com rótulo (3). Entretanto, para o caso do ponteiro de cor (painel com rótulo (2) da Figura 19(a), o painel 2 da Tabela 5), a abordagem não-linear em (VAN GINNEKEN et al., 2002) forneceu os melhores resultados. No caso do painel com ponteiro de cor, o esquema ASM proposto alcançou um melhor posto do que o método ASM (BEHAINE; SCHARCANSKI, 2012b), provavelmente porque o esquema ASM proposto utiliza imagem a cores (modelo de cor RGB) para a representação dos perfis.

Tabela 7: Desempenho comparativo do monitoramento visual automático proposto para multipainéis analógicos de medidores, ilustrado para um sistema com  $U = 3$  painéis.

Métodos ASM	Posto painel 1	Posto painel 2	Posto painel 3	$\Omega$
ASM proposto	2	2	1	1.66
ASM (COOTES et al., 1995)	4	4	4	4.00
ASM (VAN GINNEKEN et al., 2002)	3	1	3	2.33
ASM (BEHAINE; SCHARCANSKI, 2012b)	1	3	2	2.00

## 4.4 Discussão

O método do ASM proposto para reconhecimento de objetos apresentado na seção 3.1 define pesos  $a_i$  através da Equação (14), os quais refletem um poder de discriminação das feições para os marcadores. Entretanto, quando os marcadores são afastados das bordas do quadro da imagem do objeto é necessário ter um fundo ou *background* uniforme para aproveitar a seleção de ditos marcadores a partir dos valores dos pesos  $a_i$ . As vantagens

do método do ASM proposto na seção 3.1 são duas: permite uma classificação de alto desempenho como ilustrado na seção 4.1 quando utilizado em aplicações de reconhecimento de faces onde há deformações, e permite identificar os pontos mais relevantes para o reconhecimento de objetos. A desvantagem do método do ASM proposto na seção 3.1 é que afeta-se o desempenho sistêmico pelo pré-processamento (i.e. controle de luminosidade, detector de face, etc).

O método do ASM proposto para a melhora da comparação dos marcadores apresentado na seção 3.2, utiliza definições de marcadores adequados para objetos com bordas bem definidas. O método do ASM proposto na seção 3.2 pode ser utilizado para a leitura de multipainéis como detalha-se na seção 4.3. Os resultados do método ASM proposto e outros métodos ASM testados estão ilustrados na seção 4.3. Os métodos ASM testados para multipainéis foram avaliados com base na média dos posicionamentos ou postos em relação ao desempenho medido em termos da média do erro absoluto ( $\mu$ ) em cada painel, para não depender das diferentes unidades de medidas que tem cada painel. O posto é um posicionamento global que também permite que a avaliação comparativa dos métodos ASM testados não seja sensível à qualidade das marcações das mínimas escalas de divisão inseridas pelos fabricantes dos diferentes painéis, (i.e. o painel com rótulo (2) da Figura 19(a) apresenta uma escala mínima de divisão  $\check{e} = 1$  Volt AC, cujo valor não reflete a tolerância real do dito painel como ilustrado na Tabela 5 em termos de resultados de leitura humana e leitura através dos métodos ASM testados quando comparado à referência, porém adotou-se no painel 2 uma escala de divisão melhor marcada no painel  $\check{e} = 5$  Volts AC). As vantagens do método proposto na seção 3.2 utilizando os materiais da seção 4.3 para monitoramento de multipainéis são: permite uma leitura sem contato de vários tipos diferentes de painéis simultaneamente e pode ser útil para laços redundantes de monitoramento remoto. Entretanto, a desvantagem dos métodos visuais de monitoramento é que precisam de uma linha de visão entre o sensor e o painel, o que pode ser um problema no caso do painel sofrer queda ou vibrações que afastem o painel para fora da área de alcance do sensor de imagem.

## 5 CONCLUSÕES

Os modelos ativos de formas (ASMs) são modelos estatísticos de formas que procuram fazer uma comparação adaptativa das formas de um objeto. Os modelos ativos de formas têm sido utilizados frequentemente para localizar feições em imagens de objetos que podem sofrer deformações. Entretanto, o desempenho do ASM pode ser degradado substancialmente em presença de ruído ou perto dos contornos do objeto. Neste sentido, propõe-se nesta tese um novo esquema de seleção de marcadores em ASM para melhorar o seu desempenho em aplicações de reconhecimento de objetos, e um ASM modificado para a melhora da comparação de objetos em imagens a cores.

Um novo método de seleção de feições é introduzido no esquema ASM proposto para reconhecimento de objetos e permite adequar os modelos ativos de formas para reconhecimento de padrões em objetos onde há deformações. O dito método seleciona os marcadores consistentes, onde feições relevantes do objeto estão localizadas, e associa pesos altos às feições correspondentes encontradas na etapa de classificação dos objetos. Os resultados experimentais apresentados são promissores e indicam que a abordagem tem a vantagem que tende a melhorar o desempenho do ASM, levando a melhoras dos resultados finais de classificação dos objetos. A desvantagem do dito método é que afeta-se o desempenho sistêmico pelo pré-processamento (i.e. controle de luminosidade, qualidade da detecção do objeto, etc.). O método de ASM com seleção de feições também pode ser utilizado em aplicações de monitoramento visual automático de medidores com painel analógico, pois sugere quais são os marcadores adequados para uma classificação por intervalos do painel. Os resultados experimentais apresentados indicam que o método

com base em ASM é suficientemente exato para a leitura padrão de medidores em aplicações de leitura redundante. A exatidão média foi aceitável para as formas de painéis testadas. Esta técnica de medição sem contato utiliza uma tecnologia de baixo consumo de potência nos painéis, o que é adequado para sistemas de monitoramento redundante, e pode ser útil em tarefas críticas que envolvem distribuição e geração de potência, áreas radioativas e zonas químicas, onde é desejável ter componentes isolados e redundantes. A abordagem apresentada tem potencial para ser adaptada em diferentes formas de painéis, e a região de interesse (ROI) pode ser detectada automaticamente na imagem do painel.

O ASM modificado para a melhora da comparação de objetos em imagens a cores, baseia-se na definição adequada de conjuntos de marcadores que representam o objeto e na extensão de perfis de tons de cinza para os diferentes canais que representam as cores. Definem-se dois conjuntos de marcadores para um PDM, a saber, o conjunto de marcadores identificados na imagem, e o conjunto de marcadores que são interpolados a partir dos marcadores identificados. Utiliza-se um perfil de tons de cinza na vizinhança de cada uma dos marcadores detectados sobre uma imagem de entrada (i.e. imagem a cores) do objeto, para melhorar a comparação da forma. Um modelo de aparência local é criado a partir destes perfis de tons de cinza durante a etapa de treinamento de um ASM, que a sua vez, é também utilizado nas imagens de entrada do objeto. O perfil de tons de cinza auxilia na detecção dos marcadores através da estimativa das localizações mais parecidas da imagem de entrada. Este ASM modificado proposto procura melhorar a robustez da comparação e pode ser utilizado em aplicações de monitoramento visual automático de multipainéis. Supõe-se que um ASM foi treinado para ler uma dada medida com base em um conjunto de imagens de seu painel. O esquema de leitura do medidor analógico apresentado utiliza uma combinação de diferentes tipos de marcadores e uma modelagem adequada da sua deslocação ao longo da escala do medidor analógico. O esquema proposto utiliza o perfil característico em torno de cada um dos marcadores detectados em uma imagem de entrada (i.e. imagem a cores) do painel do medidor, para melhorar a comparação da forma em relação com outras alternativas ASM. Este método foi concebido para ler vários painéis analógicos remotamente sem contato e com segurança, em medidores onde

existem diferentes visores analógicos (com formas e cores distintas). Baseando-se nos resultados experimentais, o esquema ASM proposto tem potencial para um desempenho de leitura remota melhor do que as leituras humanas dos mesmos painéis analógicos. O esquema ASM modificado proposto alcançou o melhor desempenho comparativo geral quando comparado com outros esquemas ASM de leitura. Entretanto, a desvantagem dos métodos visuais de monitoramento é que precisam de uma linha de visão entre o sensor e o painel, o que pode ser um problema no caso do painel sofrer de queda ou vibrações que afastem o painel fora da área de alcance do sensor de imagem.

Os trabalhos futuros vão se concentrar na exploração de novas aplicações e em melhorar a eficiência dos métodos para atingir um desempenho em tempo real. Uma característica do método do ASM com seleção de feições é que permite uma representação eficiente em termos de informações relevante localizada, isto é, tem potencial para aplicações de reconhecimento colaborativo de objetos, onde seja necessária a transmissão eficiente de feições do objeto entre quantidades grandes de sensores de imagens.



## REFERÊNCIAS

ALEGRIA, F.; SERRA, A. Automatic Calibration of Analog and Digital Measuring Instruments Using Computer Vision. **IEEE Transactions on Instrumentation and Measurement**, Piscataway, v.49, n.1, p.94–99, 2000.

BEHAINE, C.; SCHARCANSKI, J. Enhancing the Performance of Active Shape Models in Face Recognition Applications. **IEEE Transactions on Instrumentation and Measurement**, Piscataway, v.61, n.8, p.2330–2333, 2012.

BEHAINE, C.; SCHARCANSKI, J. A New Approach for Automatic Visual Monitoring of Analog Meter Displays. In: INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC); MAY 2012 ; EUROPE, 2012, Graz. **Proceedings. . .** Graz: IEEE, 2012. p.230–233.

BREJL, M.; SONKA, M. Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples. **IEEE Transactions on Medical Imaging**, Iowa City, v.19, n.10, p.973–985, 2000.

CERROLAZA, J. et. al. Full Multiresolution Active Shape Models. **Journal of Mathematical Imaging and Vision**, New York, v.44, n.3, p.463–479, 2012.

COOTES, T. Model-Based Methods in Analysis of Biomedical Images. In: BALDOCK, R.; J.GRAHAM (Ed.). **Image Processing and Analysis**. Oxford: Oxford University Pres, 2000. p.223–248.

COOTES, T. et. al. Trainable method of parametric shape description. **Image and Vision Computing**, Chapel Hill, v.10, n.5, p.289 – 294, 1992.

COOTES, T.; EDWARDS, G.; TAYLOR, C. Active appearance models. In: BURKHARDT, H.; NEUMANN, B. (Ed.). **Computer Vision - ECCV 98**. Freiburg: Springer Berlin Heidelberg, 1998. p.484–498. Lecture Notes in Computer Science, v.1407.

COOTES, T. F. et. al. Active Shape Models - Their Training and Application. **Computer Vision and Image Understanding**, West Lafayette, v.61, n.1, p.38–59, 1995.

COOTES, T. F.; TAYLOR, C. J. Active Shape Model Search Using Local Grey-Level Models: a quantitative evaluation. In: BRITISH MACHINE VISION CONFERENCE; SEP 1993, 1993, Manchester. **Proceedings. . .** Manchester: BMVA, 1993. p.1–64.

DASH, M.; LIU, H. Feature Selection for Classification. **Intelligent Data Analysis**, Ottawa, v.1, p.131–156, 1997.

DAVIES, R. H. et. al. 3D Statistical Shape Models Using Direct Optimisation of Description Length. In: EUROPEAN CONFERENCE ON COMPUTER VISION-PART III, 7., 2002, London. **Proceedings. . .** London: Springer-Verlag, 2002. p.3–20.ECCV '02.

DEMIREL, H.; ANBARJAFARI, G. Data fusion boosted face recognition based on probability distribution functions in different colour channels. **EURASIP J. Adv. Signal Process**, New York, v.2009, p.25:1–25:10, 2009.

DENG, C.; XIAOFEI, H.; JIAWEI, H. SRDA An Efficient Algorithm for Large-Scale Discriminant Analysis. **IEEE Transactions on Knowledge and Data Engineering**, Burnaby, v.20, n.1, p.1–12, 2008.

DUDA, R.; HART, P.; STORK, D. **Pattern Classification**. New York: Wiley, 2001. 680p.

DUTA, N.; SONKA, M. Segmentation and interpretation of MR brain images: an improved active shape model. **IEEE Transactions on Medical Imaging**, Iowa City, v.17, n.6, p.1049–1062, 1998.

ESCOLANO, F. et. al. Deformable templates for tracking and analysis of intravascular ultrasound sequences. In: PELILLO, M.; HANCOCK, E. R. (Ed.). **Energy Minimization Methods in Computer Vision and Pattern Recognition**. Venice. Venice: Springer Berlin Heidelberg, 1997. p.521–534. Lecture Notes in Computer Science, v.1223.

ESSEX university face database vision group. Disponível em:  
< <http://cswww.essex.ac.uk/mv/allfaces/faces94.html> > . 2011. Acesso em: 10 jun. 2011

FIGUEIREDO, M. A. T.; JAIN, A. Unsupervised learning of finite mixture models. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, v.24, n.3, p.381–396, 2002.

FLETCHER, R.; FLETCHER, S. **Clinical Epidemiology: the essentials**. Philadelphia: Lippincott Williams & Wilkins, 2005. 288p. Epidemiology/Biostatistics.

FROBA, B. et. al. Real-time active shape models for face segmentation. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2001, Thessaloniki. **Proceedings. . .** Piscataway: IEEE, 2001. v.1, p.205–208.

GANESH, K.; THANUSHKODI, K. Design of Virtual Instrumentation for Remote Testing of Instruments using Image Processing. **GVIP Journal**, Las Palmas de Gran Canaria, v.6, n.4, p.1–6, 2006.

GEISSER, S. **Predictive Inference**. New York: Taylor & Francis, 1993. 240p. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

- HILL, A.; COOTES, T.; TAYLOR, C. Active Shape Models and the Shape Approximation Problem. **Image and Vision Computing**, Chapel Hill, v.14, n.8, p.601–607, 1996.
- HILL, A.; TAYLOR, C.; BRETT, A. A framework for automatic landmark identification using a new method of nonrigid correspondence. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, v.22, n.3, p.241–251, 2000.
- HUANG, H. I. et. al. Study on image feature selection: a genetic algorithm approach. In: IET INTERNATIONAL CONFERENCE ON FRONTIER COMPUTING. THEORY, TECHNOLOGIES AND APPLICATIONS, 2010., 2010, Taichung. **Proceedings. . .** Taichung: IET, 2010. p.169–174.
- ISARD, M.; BLAKE, A. CONDENSATION- Conditional Density Propagation for Visual Tracking. **International Journal of Computer Vision**, New York, v.29, n.1, p.5–28, 1998.
- JAIN, A.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition a review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, CA, USA, v.22, n.1, p.4–37, 2000.
- JAIN, A.; ZONGKER, D. Feature selection: evaluation, application, and small sample performance. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, v.19, n.2, p.153–158, 1997.
- JENKINS, R.; BURTON, A. M. 100% Accuracy in Automatic Face Recognition. **Science**, New York, v.319(5862), n.14, p.435, 2008.
- JUNMO, K.; MÜJDAT, Ç.; ALAN, S. W. Nonparametric shape priors for active contour-based image segmentation. **Signal Processing**, Stockholm, v.87, n.12, p.3021–3044, 2007.
- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: active contour models. **International Journal of Computer Vision**, New York, v.1, n.4, p.321–331, 1988.
- KAUFMAN, M.; SEIDMAN, A. **Handbook of Electronics Calculations For Engineers and Technicians**. 2nd. ed. New York: McGraw-Hill, 1988. 864p.
- LADHA, L.; DEEPA, T. Feature selection methods and algorithms. **International Journal on Computer Science and Engineering**, Tamil Nadu, v.3, n.5, p.1787–1797, 2011.
- LEI, S. A Feature Selection Method Based on Information Gain and Genetic Algorithm. In: INTERNATIONAL CONFERENCE ON COMPUTER SCIENCE AND ELECTRONICS ENGINEERING (ICCSEE) 2012, 2012, Hangzhou. **Proceedings. . .** Los Alamitos: IEEE, 2012. v.2, p.355–358.
- LI, H.; YEZZI, A. Local or Global Minima: flexible dual-front active contours. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, v.29, n.1, p.1–14, 2007.

- LI, Y.; FENG, J. Sparse Representation Shape Model. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP); SEP 2010, 2010, Hong Kong. **Proceedings**. . . Piscataway: IEEE, 2010. p.2733–2736.
- LIPSON, P.; et. al. Deformable templates for feature extraction from medical images. In: FAUGERAS, O. (Ed.). **Computer Vision - ECCV 90**. Antibes: Springer Berlin Heidelberg, 1990. p.413–417. Lecture Notes in Computer Science, v.427.
- LIU, B.; WU, S.; ZOU, S. Automatic Detection Technology of Surface Defects on Plastic Products Based on Machine Vision. In: MECHANIC AUTOMATION AND CONTROL ENGINEERING (MACE) ASIA AND PACIFIC, 2010, Wuhan. **Proceedings**. . . Red Hook: MACE, 2010. p.2213–2216.
- MILBORROW, S.; NICOLLS, F. Locating Facial Features with an Extended Active Shape Model. In: EUROPEAN CONFERENCE ON COMPUTER VISION: PART IV, 10., 2008, Marseille. **Proceedings**. . . Berlin: Springer-Verlag, 2008. p.504–513. ECCV'08.
- MOLINA, L.; BELANCHE, L.; NEBOT, A. Feature selection algorithms: a survey and experimental evaluation. In: IEEE INTERNATIONAL CONFERENCE ON DATA MINING, ICDM 2003, 2002, Maebashi City. **Proceedings**. . . Los Alamitos: IEEE, 2002. p.306–313.
- MORTENSEN, E. N.; BARRETT, W. A. Intelligent scissors for image composition. In: COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 22., 1995, New York. **Proceedings**. . . New York: ACM, 1995. p.191–198. SIGGRAPH '95.
- PENG, H.; LONG, F.; DING, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Los Alamitos, v.27, n.8, p.1226–1238, 2005.
- PEREZ, P.; BLAKE, A.; GANGNET, M. JetStream: probabilistic contour extraction with particles. In: EIGHTH IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, ICCV 2001, 2001, Vancouver. **Proceedings**. . . Vancouver: IEEE, 2001. v.2, p.524–531.
- PUDIL, P.; NOVOVICOVÁ, J.; KITTLER, J. Floating search methods in feature selection. **Pattern Recognition Letters**, Uppsala, v.15, n.11, p.1119 – 1125, 1994.
- REIS, E. **Estatística Descritiva**. 7. ed. Lisboa: Edições Sílabo, 2008. 248p.
- SCHAFFER, A.; GOULD, J. Visual monitoring of multi-channel displays. **IEEE Transactions on Human Factors in Electronics**, Cambridge, v.HFE-7, n.2, p.69–76, 1966.
- SHI, J.; SAMAL, A.; MARX, D. How effective are landmarks and their geometry for face recognition? **Computer Vision and Image Understanding**, West Lafayette, v.102, n.2, p.117–133, 2006.

TEDÑAN, R.; BECERRA, J.; DURO, R. J. Using classifiers as heuristics to describe local structure in Active Shape Models with small training sets. **Pattern Recognition Letters**, Uppsala, v.34, n.14, p.1710–1718, 2013.

TERZOPOULOS, D.; SZELISKI, R. Tracking with Kalman snakes. In: IN BLAKE, A.; YUILLE, A. L. (Ed.). **Active Vision**. Cambridge: MIT Press Cambridge Massachusetts, 1992. p.3–20.

TURK, M.; PENTLAND, A. Face recognition using eigenfaces. In: CVPR '91. IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 1991, Maui. **Proceedings**. . . Maui: IEEE, 1991. p.586–591.

VAN GINNEKEN, B. et. al. Active shape model segmentation with optimal features. **IEEE Transactions on Medical Imaging**, Iowa City, v.21, n.8, p.924–933, 2002.

VIGNOLO, L. et. al. An evolutionary wrapper for feature selection in face recognition applications. In: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS (SMC), 2012, Seoul. **Proceedings**. . . Red Hook: IEEE, 2012. p.1286–1290.

VINH, N. X.; EPPS, J.; BAILEY, J. Information Theoretic Measures for Clusterings Comparison: variants, properties, normalization and correction for chance. **Journal of Machine Learning Research**, Vancouver, v.11, p.2837–2854, 2010.

WAN, K.-W.; LAM, K.-M.; NG, K.-C. An accurate active shape model for facial feature extraction. **Pattern Recognition Letters**, Uppsala, v.26, n.15, p.2409–2423, 2005.

WANG, W. et. al. An Improved Active Shape Model for Face Alignment. In: FOURTH IEEE INTERNATIONAL CONFERENCE ON MULTIMODAL INTERFACES (ICMI); OCT 2002, Pittsburgh. **Proceedings**. . . Piscataway: IEEE, 2002. p.523–528.

YARI, Y.; CAVALCANTI, P.; SCHARCANSKI, J. Estimation of the head pose based on monocular images. In: INTERNATIONAL CONFERENCE OF IMAGE AND VISION COMPUTING NEW ZEALAND (IVCNZ), 25., 2010, Queenstown. **Proceedings**. . . Queenstown: IEEE, 2010. p.1–7.

YARI, Y.; SCHARCANSKI, J. An approach for fast human head pose estimation. In: SPIE 8063, MOBILE MULTIMEDIA/IMAGE PROCESSING, SECURITY, AND APPLICATIONS, 2011, Orlando. **Proceedings**. . . Bellingham: SPIE Digital Library, 2011. v.8063, p.1–7.

YUILLE, A. L.; HALLINAN, P. W.; COHEN, D. S. Feature extraction from faces using deformable templates. **International Journal of Computer Vision**, New York, v.8, n.2, p.99–111, 1992.

ZHAO, S. et. al. Research on Remote Meter Automatic Reading Based on Computer Vision. In: TRANSMISSION AND DISTRIBUTION CONFERENCE AND EXHIBITION; AUG 2005 ; ASIA AND PACIFIC, 2005, Dalian. **Proceedings**. . . Dalian: IEEE/PES, 2005. p.1–4.

ZHENG, N. et. al. 50 Years of Image Processing and Pattern Recognition in China.  
**IEEE Intelligent Systems**, Tucson, v.23, n.6, p.33–41, 2008.

ZHENG, Z. et. al. Facial feature localization based on an improved active shape model.  
**Information Sciences**, Edmonton, v.178, n.9, p.2215–2223, 2008.

## APÊNDICE A INFORMAÇÃO MÚTUA AJUSTADA

A informação mútua ajustada (AMI) é uma medida adimensional normalizada da dependência entre duas variáveis a partir de uma abordagem da entropia, cuja aplicação tem conseguido otimizar o comprimento de dados aglomerados ou *clusters* para a classificação (VINH; EPPS; BAILEY, 2010). A definição da AMI, denotada  $A$ , entre as variáveis  $u$  e  $v$ ,  $0 \leq A \leq 1$ , é:

$$A = \frac{MI(u, v) - E\{MI(u, v)\}}{\max\{\hat{H}(u), \hat{H}(v)\} - E\{MI(u, v)\}} \quad (32)$$

onde  $MI(u, v)$ ,  $E\{MI(u, v)\}$  denota a informação mútua entre as variáveis  $u$  e  $v$ , e o seu valor esperado, respectivamente;  $\hat{H}(u)$  e  $\hat{H}(v)$  são as entropias do  $u$  e  $v$ .

Observa-se na Equação (32) que o valor especial do ajuste é a expressão do valor esperado da informação mútua,

$$E\{MI(u, v)\} = \sum_{i=1}^Z \sum_{n=1}^C \sum_{\varpi_{i,n}=\max(d_i+b_n-\aleph, 0)}^{\min(d_i, b_n)} \frac{\varpi_{in}}{\aleph} \log \left( \frac{\aleph \varpi_{in}}{d_i b_n} \right) \varphi_{in} \quad (33)$$

onde  $Z$ ,  $C$  é número de elementos das variáveis  $u$  e  $v$  respectivamente. O termo que atribui os pesos às ocorrências é dado por

$$\varphi_{in} = \frac{d_i! b_n! (\aleph - d_i)! (\aleph - b_n)!}{\aleph! \varpi_{in}! (d_i - \varpi_{in})! (b_n - \varpi_{in})! (\aleph - d_i - b_n + \varpi_{in})!} \quad (34)$$

onde  $\varpi_{in}$  são os termos da matriz de ocorrência  $\Theta = [\varpi_{in}]$ , cuja estatística em termos de  $\aleph = \sum_{in} \varpi_{in}$ ,  $d_i$  e  $b_n$ , com  $i = 1, \dots, Z$  e  $n = 1, \dots, C$  ilustra-se na Tabela 8.

A Tabela 8 facilita o cálculo da  $MI(u, v)$ ,  $\hat{H}(u)$  e  $\hat{H}(v)$ , cujas expressões são dadas

Tabela 8: Estatística dos termos da matriz de ocorrências  $\Theta = [\varpi_{in}]$ , com  $\varpi_{in} = |u_i \cap v_n|$  para as variáveis  $u$  e  $v$ .

$u \setminus v$	$v_1$	$v_2$	$\dots$	$v_C$	$\Sigma$
$u_1$	$\varpi_{11}$	$\varpi_{12}$	$\dots$	$\varpi_{1C}$	$d_1$
$u_2$	$\varpi_{21}$	$\varpi_{22}$	$\dots$	$\varpi_{2C}$	$d_2$
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$
$u_Z$	$\varpi_{Z1}$	$\varpi_{Z2}$	$\dots$	$\varpi_{ZC}$	$d_Z$
$\Sigma$	$b_1$	$b_2$	$\dots$	$b_C$	$\aleph = \sum_{in} \varpi_{in}$

como a seguir:

$$MI(u, v) = \sum_{i=1}^Z \sum_{n=1}^C \frac{\varpi_{in}}{\aleph} \log \left( \frac{\varpi_{in}/\aleph}{d_i b_n / \aleph^2} \right) \quad (35)$$

$$\hat{H}(u) = - \sum_{i=1}^Z \frac{d_i}{\aleph} \log \left( \frac{d_i}{\aleph} \right) \quad (36)$$

$$\hat{H}(v) = - \sum_{n=1}^C \frac{b_n}{\aleph} \log \left( \frac{b_n}{\aleph} \right) \quad (37)$$



## APÊNDICE B ESTIMATIVA DO $\alpha$ ÓTIMO

O  $\alpha$  ótimo denotado  $\alpha_o$ , associado ao valor máximo de  $PPV$ , foi estimado utilizando um procedimento de validação cruzada nos dados de treinamento. Para cada valor  $0 < \alpha \leq 1$ , tem-se  $JK$  validações em cada dimensionalidade  $R$ . Selecionam-se  $(J - 1)$  imagens por classe para o treinamento e uma imagem por classe para a validação. As imagens de treinamento e validação são mutuamente exclusivas. No algoritmo 1 ilustram-se os detalhes do procedimento da estimativa do  $\alpha$  ótimo.

---

### Algorithm 1 Estimativa do $\alpha$ ótimo utilizando validação cruzada

---

```

1: procedure [ $\alpha_o$ ] = VALIDAÇÃOCRUZADA( $\mathbf{v}_\alpha(0, 1]$ ,  $\mathbf{r}_R[1, N]$ ,  $K$ )
2:   Inicializar os subconjuntos de treinamento  $\{T_{J-1}\}$  ▷  $J$  arranjos
3:   Inicializar os subconjuntos de validação  $\{V_1\}$  ▷  $J$  arranjos
4:   Inicializar  $\{T_{J-1}, V_1, \}$  ▷ Mutuamente exclusivos
5:   for all  $\alpha \in v_\alpha(0, 1]$  do ▷ Para as amostras do  $\mathbf{v}_\alpha$ 
6:     for  $j = 1 : J$  do ▷ Para todos os arranjos
7:       Selecione  $\{T_{J-1}\}_j$  ▷ Treinamento
8:       Calcule  $a_i$  ▷ Equação (14)
9:       for all  $k \in \{K\}$  do ▷ Treina o modelo para as classes
10:        Calcule  $\Omega_{ck}$  ▷ Equação (15)
11:      end for
12:      Selecione  $\{T_1\}_j$  ▷ Validação
13:      for all  $R \in r_R[1, N]$  do ▷ Para as amostras do  $\mathbf{r}_R$ , com  $R \leq N$ 
14:        Calcule  $k'$  ▷ Equação (16)
15:        Calcule  $PPV(\alpha, j, R)$ 
16:      end for
17:    end for
18:  end for
19:  Calcule  $\alpha_o = \arg \max_{\alpha \in v_\alpha(0, 1]} \{PPV(\alpha, j, R)\}$ 
20: end procedure

```

---