

## Segmentação e Restauração Digital para Eliminação de Artefatos em Imagens Frontais de Face

André Sobiecki<sup>1</sup>, Luiz A. P. Neves<sup>2</sup>, Gilson A. Giraldi<sup>3</sup>, Gilka J. F. Gattás<sup>4</sup> e Carlos E. Thomaz<sup>1</sup>

<sup>1</sup>FEI, São Bernardo do Campo, São Paulo, Brasil

<sup>2</sup>UFPR, Curitiba, Paraná, Brasil

<sup>3</sup>LNCC, Petrópolis, Rio de Janeiro, Brasil

<sup>4</sup>USP, São Paulo, São Paulo, Brasil

### Abstract

*There is a variety of applications nowadays that has used the face image as a relevant information for recognizing subjects. However, depending on the application, such as in the problem of identifying missing people based on paper archiving, the face images can become noisy for digitalization due to its low original resolution, poor quality of the photographic paper and the presence of other artifacts like folds and stamps. In practice, all these artifacts make the process of automatically recognizing faces even harder. Therefore, we propose in this work a computational framework for segmentation and automatic restoration of image artifacts based on a statistical model built from 200 frontal faces with neutral expression. The effectiveness of the framework has been compared with other methods and shows promising results.*

### 1. Introdução

O reconhecimento facial é o processo de identificação de uma pessoa com base em informações obtidas a partir da aparência facial [3]. O desempenho dos algoritmos de reconhecimento facial depende da qualidade do pré-processamento da imagem do rosto, da precisão da representação facial e da eficiência do classificador [7]. Por isso, é de fundamental importância que a imagem a ser utilizada em um trabalho de reconhecimento facial apresente apenas as características faciais discriminantes com o mínimo de artefatos possível.

Há aplicações que utilizam a face como informação relevante para identificação de pessoas. Por exemplo, o site do governo federal brasileiro de crianças e adolescentes desaparecidos [9] disponibiliza publicamente imagens faciais de pessoas sendo que muitas dessas fotos são antigas e possuem rasuras como dobraduras, arranhões, luminância irregular, bolor, carimbos e escritos diversos sobre a imagem digitalizada. Na prática, essas rasuras dificultam o reconhecimento facial automático. Uma rasura sobre uma imagem facial pode estar em diversas cores, tons, formatos e tamanhos. Informações

prévias sobre a localização de características faciais como nariz, olhos, boca e cabelo auxiliam na identificação de rasuras.

Este trabalho propõe um arcabouço computacional de segmentação e restauração automática de rasuras baseado em um modelo estatístico construído a partir de 200 imagens faciais frontais. A localização das características faciais é feita por meio de uma imagem de referência gerada a partir de uma população amostral que fornece informações *a priori* sobre tons e localização espacial das características. As rasuras identificadas na imagem são posteriormente restauradas por um método de *inpainting* [13, 15]. Os experimentos deste estudo mostram que a segmentação baseada no método estatístico apresenta melhores resultados de correlação com a imagem de referência do que outros métodos implementados para comparação, como os métodos baseados em entropia e no filtro de Sobel.

### 2. Arcabouço Computacional

Um modelo deve cumprir os critérios de generalidade e especificidade, onde modelos gerais são aqueles que representam as fontes de variação na aparência de imagens faciais e os modelos específicos são aqueles que limitam a variabilidade permitida de modo que apenas exemplos que cumpram os devidos critérios possam ser gerados [6].

A Figura 1 representa a seqüência de métodos proposta neste artigo para identificação e eliminação de rasuras automaticamente.



Figura 1: Arcabouço computacional de eliminação de rasuras.

Todas as imagens foram antes equalizadas [2] e pré-processadas utilizando-se um filtro 3x3 de mediana para eliminar ruídos. O filtro de mediana ordena os pixel em ordem crescente e escolhe o valor mediano para igualar aos demais pixels. As etapas de identificação de artefatos, aplicação do molde e restauração digital são descritas em detalhes nas seções seguintes.

### 3. Identificação de Artefatos

Para identificação de artefatos, são usados três métodos de segmentação automática: o método de decisão estatística proposto neste trabalho; e mais dois métodos para fazer comparações baseados em entropia [17] e no filtro de Sobel. O filtro de Sobel é bem conhecido na literatura e não será descrito. O leitor pode consultar, por exemplo, as referências [5, 11] para mais detalhes sobre esse filtro.

#### 3.1. Decisão Estatística

O método estatístico é baseado na teoria de inferência estatística [8, 10], onde amostras populacionais podem gerar informações *a priori* e, a partir destas informações, é possível tomar decisões dado um nível  $\alpha$  de significância estatística desejado.

A partir de uma população amostral de 200 imagens frontais com expressão facial neutra e normalizadas espacialmente [2], calcula-se a imagem média dessas amostras através da seguinte equação:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (1)$$

onde  $x_i$  é o vetor  $n$ -dimensional que representa a concatenação de todos os pixels da imagem  $i$  de face, e  $N$  o número total de amostras. Neste trabalho,  $N = 200$ .

Para verificar quão válida seria a afirmação de que a média amostral descrita pela equação (1) se aproximaria da média da população de faces frontais com expressão facial neutra, calculou-se o intervalo de 99% de confiança (IC) dessa estimativa, ou seja,

$$99\%IC = \bar{x} \pm 2.58 \frac{s}{\sqrt{N}}, \quad (2)$$

onde  $s$  é o desvio padrão das amostras, cuja variância  $s^2$  pode ser definida como:

$$s^2 = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2. \quad (3)$$

A Figura 2 ilustra as imagens que correspondem aos limites inferior e superior do intervalo de confiança descritos na equação (2). Observa-se que visualmente as imagens são muito parecidas, pois possuem correlação de 0,998942 e 0,998962 com a imagem média amostral.

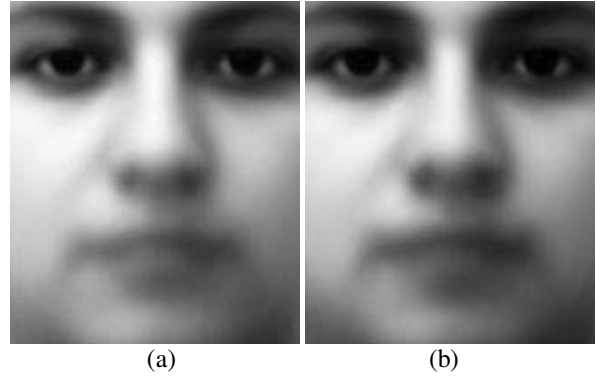


Figura 2: Intervalo de confiança da imagem média: (a) Limite inferior; (b) Limite superior.

Portanto, considerando estatisticamente válida a estimativa amostral da média, a decisão sobre a identificação de artefatos está baseada simplesmente na definição da região crítica das diferenças significantes. Em outras palavras, calcula-se o valor  $z$  da diferença entre uma imagem de face  $x$  e a imagem da média amostral  $\bar{x}$ , supondo que a distribuição das tonalidades dos pixels segue uma densidade de probabilidade Gaussiana com média nula e variância 1, isto é:

$$z = \frac{x - \bar{x}}{s}. \quad (4)$$

Quanto maior for o valor absoluto de  $z$  maior será a significância estatística dessa diferença, pixel a pixel. A Figura 3 mostra a imagem média e o desvio padrão de todos os pixels utilizados na equação (4) para a identificação das rasuras. Para os experimentos desse trabalho, o valor do nível de significância  $\alpha$  foi fixado em 5%, que equivale a um valor  $z$  padrão de 1,96 [8].



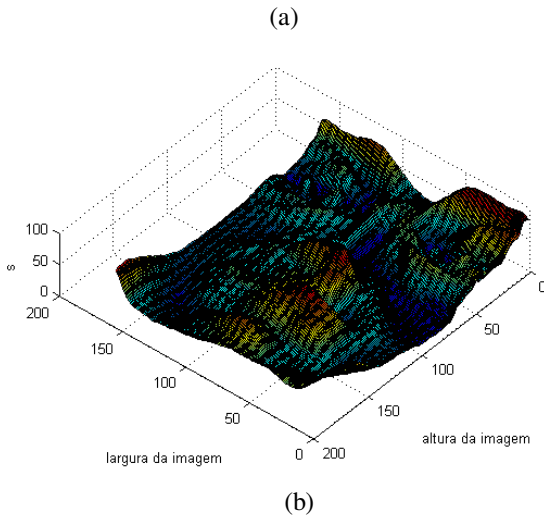


Figura 3: Medidas estatísticas de referência: (a) Imagem da média amostral; (b) Desvio padrão de todos os pixels.

### 3.2. Entropia

Em processamento de imagens, uma forma de distinguir o plano de fundo da imagem do objeto de interesse, é através da aplicação de um limiar. Uma forma automática para encontrar este limiar é via a otimização da entropia, a qual fornece uma medida da quantidade de informação na imagem. Neste trabalho, considera-se duas definições para entropia: entropia de Shannon e de Tsallis [17].

Seja então um sistema com  $n$  estados possíveis, aos quais estão associadas probabilidades  $(p_1, \dots, p_n)$ , onde  $p_i \in [0,1]$  para todos  $i = 1, \dots, n$  e  $\sum p_i = 1$ . Shannon mediu a quantidade de informação do sistema pela expressão [16, 17, 18, 19, 20]:

$$S = - \sum_i p_i \ln(p_i) . \tag{5}$$

Por outro lado, Tsallis propôs a seguinte medida de informação:

$$S_q(p_1, \dots, p_k) = \frac{1 - \sum_{i=1}^k p_i^q}{q - 1} . \tag{6}$$

$p_i$  é a probabilidade de ocorrência do nível de cinza  $i$  (frequência relativa) na imagem ( $0 \leq i \leq 255$ ). Cada limiar  $t$  divide a imagem e, conseqüentemente, o histograma da mesma em duas regiões, por exemplo A e B:  $p_1, p_2, \dots, p_t$  e  $p_{t+1}, p_{t+2}, \dots, p_{255}$ . Assim, pode-se normalizar as distribuições resultantes para obter as probabilidades associadas às regiões A e B e, conseqüentemente, computar as entropias  $S^A$  e  $S^B$  pela expressão (5) ou  $S_q^A$  e  $S_q^B$  pela expressão (6). Pode-se mostrar que [21]:

$$S^{A+B}(t) = S^A + S^B , \tag{7}$$

$$S_q^{A+B}(t) = S_q^A + S_q^B + (1-q)S_q^A S_q^B , \tag{8}$$

onde A+B representa o sistema (imagem) composto pelos subsistemas A e B. A idéia básica da aplicação do conceito de entropia para segmentação é que o limiar ótimo é aquele que maximiza  $S^{A+B}$  (ou  $S_q^{A+B}$ ), ou seja:

$$T = \max(S^{A+B}(t)) . \tag{9}$$

Para entropia de Tsallis, o parâmetro  $q$  é mais um grau de liberdade na teoria que foi determinado experimentalmente através de um teste em uma imagem com o  $q$  variando de 0.1 á 0.9. A Figura 4 mostra um exemplo de segmentação por esses dois métodos de entropia descritos brevemente nesta sub-seção.

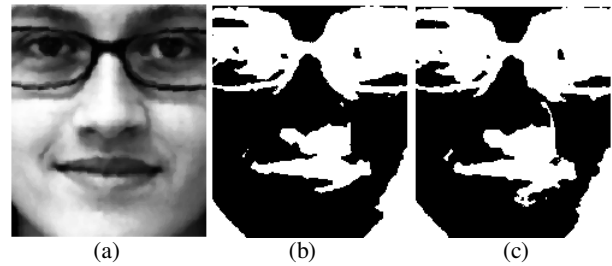


Figura 4: Segmentação por entropia: (a) Imagem original; (b) Imagem segmentada pela entropia de Shannon; (c) Imagem segmentada pela entropia de Tsallis.

### 3.3. Morfologia Matemática

Para corrigir as imperfeições das segmentações baseadas na decisão estatística e nos métodos de entropia, finaliza-se a etapa de identificação de artefatos do arcabouço computacional com a utilização de operadores de morfologia matemática.

Basicamente, a morfologia matemática é um método para extração de regiões de interesse em uma imagem (binária ou em tons de cinza), tornando possível representar e descrever melhor a forma de uma região através da detecção e suavização das suas respectivas bordas [1, 4, 11].

Foi suficiente a aplicação do operador morfológico de dilatação [1] com elemento estruturante 3x3 em uma única iteração para recuperar ou aumentar o tamanho da rasura e garantir que o método de restauração digital considerasse da forma mais abrangente possível a região rasurada existente.

## 4. Aplicação do Molde

Como há regiões faciais que podem perder suas características naturais devido ao processo de restauração digital adotado, aplica-se um molde binário para cobrir as regiões dos olhos, do nariz e da boca, conforme Figura 5. Os pixels ativos (brancos) representam as regiões a serem desconsideradas e os pixels pretos representam as regiões que podem sofrer restauração.



Figura 5: Molde usado para desconsiderar regiões que devem manter a sua originalidade.

A criação deste molde foi possível porque as imagens usadas neste trabalho são normalizadas espacialmente [2], então tem-se como informação *a priori* a localização dos olhos, do nariz e da boca.

## 5. Restauração Digital

Após a segmentação automática das rasuras, é utilizado um método de restauração digital conhecido como *inpainting* [13, 15], onde a reconstrução de um pixel é baseada na assimilação das informações contidas nos seus pixels vizinhos.

Basicamente, o *inpainting* digital pode ser entendido como uma técnica de restauração de imagens ou vídeos danificados por meio de interpolação de imagens. A técnica pode ser utilizada na restauração de fotos, em efeitos especiais, remoção de objetos, recuperação de margens e difusões de curvatura orientada [14]. O método *inpainting* busca reconstruir as características básicas da imagem e preencher dados desconhecidos em uma região de interesse [13]. Neste trabalho, essa técnica utiliza a imagem molde ilustrada na Figura 5 para determinar a localização espacial das regiões faciais adequadas para processamento do *inpainting*.

Matematicamente, pode-se definir esse processamento numérico de propagação de informações na região de *inpainting*  $\Omega$  pela seguinte equação [15, 22]:

$$I^{t+1}(i, j) = I^t(i, j) + \Delta_k I_k^t(i, j), \quad \forall (i, j) \in \Omega \quad (10)$$

onde  $I^t(i, j)$  representa cada pixel da imagem na região de *inpainting*  $\Omega$ , na iteração  $t$ , com  $I_k^t(i, j)$  representando as atualizações dos pixels de saída para uma taxa de melhoramento  $\Delta_k$ . Ou seja, a cada passo a imagem  $I^t(i, j)$  é processada, gerando uma versão melhorada desta em  $I^{t+1}(i, j)$  [22].

## 6. Experimentos

Foram testadas 30 imagens de 3 bases distintas (10 imagens por base): (a) Imagens de trabalhos anteriores [2, 23, 24]; (b) Imagens disponibilizadas publicamente pelo *site* de crianças e adolescentes desaparecidos no Brasil [9]; (c) Imagens de voluntários que concordaram em tirar uma foto ou em emprestar uma imagem de face para digitalização [23, 24]. Há rasuras como óculos, cabelo sobre a face, carimbos e arranhões.

Para avaliar quantitativamente os resultados, calculou-se o valor da correlação das 30 imagens originais e restauradas com a imagem da média amostral ilustrada anteriormente na Figura 3(a):

$$\rho_{ij} = \frac{\sum_{k=1}^n (i_k - \bar{i})(j_k - \bar{j})}{\sqrt{\sum_{k=1}^n (i_k - \bar{i})^2} \sqrt{\sum_{k=1}^n (j_k - \bar{j})^2}}, \quad (11)$$

onde  $i$  e  $j$  são imagens, e  $i_k$  e  $j_k$  representam o pixel  $k$  dessas imagens e  $\bar{i}$  e  $\bar{j}$  são os pixels médios de cada imagem. O  $n$  tem o valor correspondente a largura de 162 multiplicada pela altura de 193 de cada imagem, que no caso desse trabalho é igual então a 31266. Para determinar o parâmetro  $q$  da entropia de Tsallis, variou-se esse parâmetro de 0.1 até 0.9 e obteve-se o melhor valor de correlação para  $q = 0.3$ .

Foi utilizado o método da correlação para avaliar os resultados ao invés do método de pico do sinal de ruído (PSNR – Peak Signal to Noise Ratio) que é um método de avaliação de restauração de imagem mais comum, porque neste trabalho não há somente ruídos que afetam a qualidade de nitidez e resolução da imagem [25]. Imagens de face contendo artefatos como carimbos, mesmo estando em boa qualidade digital, não estão em condições de serem interpretadas por sistemas de reconhecimento facial.

## 7. Resultados

A Figura 6 descreve os resultados de correlação das imagens originais e restauradas com a imagem de referência amostral. Pode-se observar que a maioria das imagens restauradas pelo método estatístico teve resultados de correlação maior com o padrão frontal de faces especificado do que as imagens originais, restauradas pelos métodos de entropia e pelo filtro Sobel.

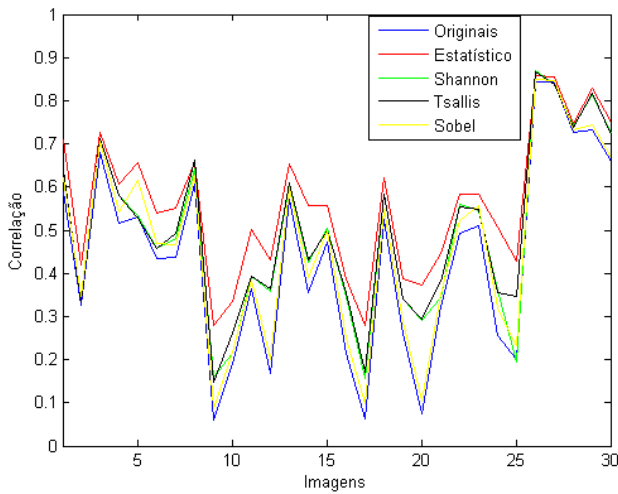


Figura 6: Correlação das 30 imagens.

As médias aritméticas dessas correlações foram: 0.4345 (originais); 0.5606 (método estatístico); 0.4915 (entropia de Shannon) e 0.5007 (entropia de Tsallis); e 0.4627 (Sobel).

A Figura 7 mostra alguns exemplos de restauração digital realizados. Pode-se perceber que o processo de restauração digital baseado no método estatístico de segmentação permitiu uma eliminação de artefatos das imagens mais conservadora e possivelmente mais realista do que os demais métodos testados.

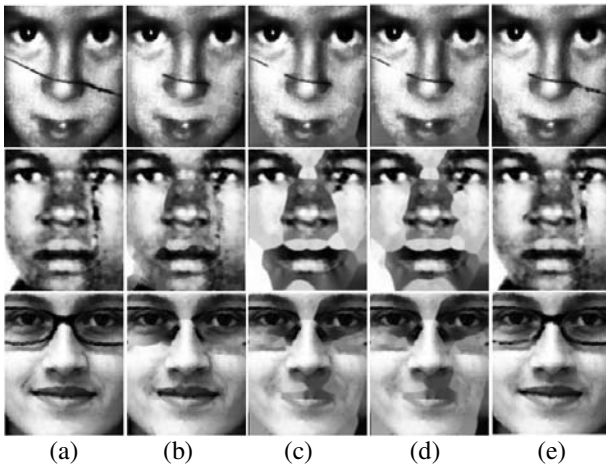


Figura 7: Alguns resultados de restauração digital: (a) imagens originais, (b) imagens segmentadas pelo método estatístico, (c) imagens segmentadas pela entropia de Shannon, (d) imagens segmentadas pela entropia de Tsallis, (e) imagens segmentadas pelo método Sobel.

## 8. Conclusão

Imagens segmentadas pelo modelo estatístico obtiveram valores maiores de correlação com a imagem de referência, considerando apenas faces frontais e de expressão facial neutra, do que as imagens originais, segmentadas pela entropia e segmentadas pelo método Sobel. No modelo estatístico, cada imagem com artefatos foi comparada com a imagem de referência para um nível de significância estatístico equivalente a 95% de certeza com relação aos defeitos na imagem. Para resultados menos conservadores, pode-se assumir valores de confiança estatística menores. No entanto, nesta situação mais regiões serão necessariamente identificadas como críticas ou, no caso, como rasuras, podendo descaracterizar a singularidade de cada imagem de face.

Devido à diversidade biométrica facial da população brasileira e considerando os resultados promissores deste estudo, pode-se propor como trabalhos futuros a criação de modelos estatísticos de face que sejam mais específicos para as diferenças de raça, idade e gênero inerentes, contribuindo principalmente para uma melhor segmentação automática de rasuras em imagens frontais de face arquivadas em papel como, por exemplo, no problema de identificação de pessoas desaparecidas. Adicionalmente, seria válido a implementação de um método que caracterizasse as regiões faciais que devem manter a originalidade como a boca, os olhos e o nariz, para os casos onde a restauração digital é feita a partir de um método de *inpainting*. As entropias de Shannon e Tsallis são métodos que medem a quantidade de informação probabilística. Através destas entropias é possível segmentar automaticamente regiões de interesse, sendo mais eficiente do que a utilização de um molde construído manualmente.

## Referências

- [1] E. Aptoula, S. Lefèvre, “A comparative study on multivariate mathematical morphology”, *Pattern Recognition, Volume 40*, November 2007, pp. 2914-2929.
- [2] V. Amaral, C. Figaro-Garcia, G. J. F. Gattas e C. E. Thomaz. Normalização espacial de imagens frontais de face em ambientes controlados e não-controlados, *Periodico Científico Eletrônico da FATEC São Caetano do Sul (FaSci-Tech)*, vol. 1, no. 1, Outubro 2009.
- [3] O. Ayinde e Y.H. Yang, “Face recognition approach based on rank correlation of Gabor-filtered images”, *Pattern Recognition, Volume 35*, 2002, pp. 1275-1289.

- [4] J.D.L. Santiago, A. Gamino, J. Salgado, V. Trujillo e A. Ortiz, “Operadores k-estadísticos para morfología matemática de conjuntos”, *Rev.fac.ing.univ. Antioquia* [online] n.48, ISSN 0120-6230, 2009, pp. 216-227.
- [5] C. Pradabpet, N. Ravinu, S. Chivapreecha, B. Knobob e K. Dehhan, “An Efficient Filter Structure for Multiplierless Sobel Edge Detection”, *Innovative Technologies in Intelligent Systems and Industrial Applications*, 2009, pp. 40-44.
- [6] G. J. Edwards, A. Lanitis, C. J. Taylor e T. F. Cootes, “Statistical models of face images — improving specificity”, *Image and Vision Computing, Volume 16, Issue 3*, 16 de marco de 1998, pp. 203-211.
- [7] B. Jun, J. Lee, e D. Kim, “A novel illumination-robust face recognition using statistical and non-statistical method”, *Pattern Recognition Letters*, 30 September 2010, pp. 1-8.
- [8] M.R. Spiegel, L.R. Stephens, *Statistics, Schaum’s Outlines*, Lisboa, 2008.
- [9] Site de “Crianças e Adolescentes Desaparecidos” da Secretaria Especial dos Direitos Humanos da Presidência da República do governo brasileiro:  
(<http://www.desaparecidos.mj.gov.br/Desaparecidos/>)
- [10] W.O. Bussab, P.A. Morettin, *Estadística Básica*, Editora Saraiva, 5ª edição, São Paulo, 2002.
- [11] R.C. Gonzalez, R.E. Woods, *Processamento de Imagens Digitais*, Editora Edgard Blücher Ltda, São Paulo, 2000.
- [13] C.A.Z. Barcelos e M.A. Batista; Image restoration using digital inpainting and noise removal; *Image and Vision Computing, Volume 25, Issue 1, January 2007, Pages 61-69*
- [14] T.K. Shih e R.C. Chang; Digital inpainting – survey and multilayer image inpainting algorithms; *Information Technology and Applications, ICITA 2005*, Page(s): 15 - 24 vol.1.
- [15] M. Bertalmio, G. Sapiro, V. Caselles e C. Ballester; Image Inpainting; *Proc. of SIGGRAPH 2000*, Page(s): 417 - 424.
- [16] M. Portes de Albuquerque, I.A. Esquef e A.R. Gesualdi Mello “Image Thresholding using Tsallis entropy”, *Pattern Recognition Letters*, Science Direct, 2004, pp. 1-7.
- [17] I.A. Esquef, “Técnicas de Entropia em Processamento de Imagens”, Tese de Mestrado, Rio de Janeiro, 2002, pp. 1-155.
- [18] S. Tong, A. Bezerianos, Y. Zhu, R. Geocadin, D. Hanley, e N. Thakor, “Monitoring Brain Injury with Tsallis Entropy”, *IEEE 23° Annual EMBS International Conference*, Instanbul, Turquia, 2001, pp. 1-3.
- [19] K. Tadaki, “The Tsallis entropy and the Shannon entropy of a universal propability”, *ISIT*, Canadá, 6-11 de Julho 2008, pp. 1-5.
- [20] M. Dobre, e L. Bolle, “Tsallis Generalized Entropy and Maximum Entropy Formalism for Spray Drops Size Distribution Function Prediction”, *Journal*, ILASS, Europa, 2001, pp. 1-1.
- [21] W. Shitong, e F.L. Chung, “Note on the equivalence relationship between Renyi-entropy based and Tsallis-entropy based image thresholding”, *Pattern Recognition Letters*, Elsevier, 2005, pp. 1-4.
- [22] U.A. Ignácio, “Aplicação de Wavelets em Inpainting Digital”, *Dissertação de Mestrado em Computação Aplicada*, 2007, pp. 1-89.
- [23] A. Sobiecki, C. E. Thomaz e L. A. P. Neves. Uma Proposta de Melhoria Automática em Imagens Faciais. *Proceedings of the V Workshop de Visao Computacional*, 5pp., Universidade Presbiteriana Mackenzie, Sao Paulo, Brazil, 2nd - 4th September 2009.
- [24] A. Sobiecki, C. E. Thomaz and L. A. P. Neves. To a better digitalization and visualization of frontal face photographs. In *proceedings of IV European Conference on Computational Mechanics, Image Processing and Visualization Symposium*, id. 1131, p. 32, Paris, France, 16th-21st May 2010.
- [25] A.N. Zamani, M.K. Awang, N. Omar e S.A. Nazeer. Image Quality Assessments and Restoration for Face Detection and Recognition System Images. *Second Asia International Conference on Modelling & Simulation*, 2008.