

Clusterização Baseada na φ -Divergência Aplicada à Segmentação de Imagens

Joniel B. Barreto e Rui F. Vigelis

Resumo— O presente trabalho apresenta uma φ -divergência, que é uma generalização das entropias relativas de Shannon e Tsallis, como medida de dissimilaridade na segmentação de imagens por clusterização. Os testes são realizados segmentando a região de texto de imagens digitalizadas com ruídos do banco de dados NoisyOffice. Baseado nos resultados obtidos, o método de clusterização proposto se mostrou mais aplicável que métodos estabelecidos, como limiar de Otsu e o K-means clássico.

Palavras-Chave— φ -Divergência, Clusterização, Segmentação.

Abstract— This work presents a φ -divergence, which is a generalization of Shannon and Tsallis relative entropies, as dissimilarity measure in image segmentation by clustering. The tests are performed by segmenting the text region of scanned images with noise from the NoisyOffice database. Based on the results, the proposed clustering method was more applicable than established methods, such as Otsu threshold and classical K-means.

Keywords— φ -Divergence, Clustering, Segmentation.

I. INTRODUÇÃO

Dentre as técnicas fundamentais em visão computacional, a segmentação, que consiste em definir regiões de interesse de uma imagem para análises posteriores, é considerada uma das operações mais importantes, pois um erro nesta etapa é refletido nas etapas seguintes. A segmentação de imagens é um processo empírico e adaptativo, que depende das particularidades da imagem. Deste modo, não há um modelo formal [1]. Dentre as técnicas utilizadas para segmentar imagem as mais comuns são a limiarização e a clusterização [2].

A limiarização fundamenta-se em converter uma imagem em tons de cinza em uma imagem binária, separando valores acima e abaixo de um valor específico, assim dividindo a imagem em duas regiões. A limiarização de Otsu [3] é uma das técnicas de limiarização mais utilizadas. Nesse método calcula-se um valor de dispersão para os níveis de intensidade dos pixels de cada lado do limiar.

A clusterização é um método de classificação não supervisionada que agrupa objetos semelhantes dentro de um mesmo grupo por meio de uma medida de similaridade. Diferentemente da limiarização, o agrupamento pode ser realizado em mais de duas regiões. Nesta técnica, primeiramente é definido o número de clusters, em seguida é localizado o centro desses clusters. Calcula-se a distância de cada ponto até o centro do cluster. Por fim, cada ponto é atribuído ao clusters que apresentar menor distância.

Joniel B. Barreto, Programa de Pós-Graduação em Engenharia Elétrica e da Computação - PPGEEC, Universidade Federal do Ceará, Campus Sobral, e-mail: joniel.bb@gmail.com; Rui F. Vigelis, Engenharia da Computação, Universidade Federal do Ceará, Campus Sobral, e-mail: rfvigelis@gmail.com

Uma das técnicas de clusterização mais conhecidas é o K-means, que é um caso especial do algoritmo de clusterização *hard* de Bregman para distância euclidiana [4]. Outros exemplos são os algoritmo LBG [5], para a distância Itakura-Saito, e o algoritmo de *information theoretic clustering* [6], para divergência Kullback-Leibler (KL). O objetivo deste trabalho é apresentar um caminho alternativo de segmentar imagens por agrupamento utilizando uma diferente medida de semelhança. Neste artigo é proposto o emprego de uma entropia relativa generalizada, obtida a partir do uso da exponencial deformada φ , para o cálculo da dissimilaridade na clusterização *hard* de Bregman.

Na literatura há vários trabalhos que analisam a utilidade de divergências de Bregman em algoritmos de clusterização. Em projetos relacionados a segmentação por clusterização comumente é utilizado a técnica K-means. Katkar e Baraskar [14] realizaram segmentação de imagens médicas usando PCA e o algoritmo K-means para extrair regiões significativas de maneira eficiente, enquanto Shmmala e Ashour [15] utilizaram três diferentes versões do K-means para segmentar imagens biológicas de tecidos e células do sangue e imagens coloridas comuns.

Zhao, Li e Zhao [16] propuseram um novo algoritmo de clusterização *fuzzy* para segmentação de imagens utilizando a distância de Mahalanobis como medida de dissimilaridade. Apresentando também as suas vantagens em relação ao algoritmo convencional *fuzzy C-means* que utiliza a distância euclidiana.

Wei, Chifn e Guo [17] apresentaram um método *Fuzzy Cluster Ensembler* baseada na divergência KL (FCE_{KL}) para segmentação de imagens. Este método é similar ao *fuzzy C-means*, diferenciando-se apenas em relação à utilização da divergência KL para realizar as associações. Gupta et al. [18] desenvolveram um trabalho que segmenta múltiplos níveis de despigmentação em imagens de vitiligo por meio de uma abordagem hierárquica de clusterização aglomerativa baseada na divergência KL simétrica e compara com outros métodos estabelecidos para trazer possíveis vantagens e desvantagens do método proposto.

O restante do artigo é organizado como segue. Na Seção II é discutido informações importantes e úteis sobre entropia relativa para o desenvolvimento do trabalho. A Seção III aborda sobre o método de clusterização proposto e apresenta o algoritmo de clusterização *hard* de Bregman. Observa-se na Seção IV a como é realizado a implementação do método proposto, assim como os resultados obtidos na segmentação. Finalmente a conclusão é relatada na Seção V.

II. ENTROPIA RELATIVA

O conceito de entropia como medida de informação foi primeiramente apresentado por Shannon [7] com a ideia de que em que cada distribuição de probabilidade $p = p_i$ é possível medir o grau de incerteza de uma variável aleatória por meio da quantidade $H(p) = -\sum_i p_i \ln p_i$. Posteriormente, Tsallis [8] definiu a expressão $H_q(p) = \frac{1}{q-1}(1 - \sum_{i=1}^n p_i^q)$ que é considerada uma generalização da entropia de Shannon. No limite de $q \rightarrow 1$, a entropia de Tsallis se aproxima da entropia de Shannon.

Com base nisso, Kullback e Leibler [9] definiram a entropia relativa de Shannon, também conhecida como divergência de Kullback-Leibler (KL), como uma medida de distância estatística entre duas distribuições de probabilidade. Ou seja, diz o quão bem uma distribuição de probabilidade q se aproxima de uma distribuição p . A divergência de Kullback-Leibler é definida pela expressão

$$D(p||q) = \sum_{i=1}^n p_i \ln \left(\frac{p_i}{q_i} \right). \quad (1)$$

Outro exemplo bem conhecido é a entropia relativa de Tsallis [10] que é obtida quando na divergência de Kullback-Leibler, substitui-se o logaritmo natural pelo q -logaritmo. Define-se a entropia relativa de Tsallis por:

$$D_q(p||q) = -\sum_{i=1}^n p_i \ln_q \left(\frac{q_i}{p_i} \right). \quad (2)$$

A. Entropia Relativa Generalizada

Vigelis e Cavalcante [11], [12], [13] observaram a similaridade entre as entropias relativas de Shannon e Tsallis, podendo ser reescritas como:

$$D(p||q) = \sum_{i=1}^n \frac{\ln(p_i) - \ln(q_i)}{\ln'(p_i)} \quad (3)$$

e

$$D_q(p||q) = \sum_{i=1}^n \frac{\ln_q(p_i) - \ln_q(q_i)}{\ln'_q(p_i)}, \quad (4)$$

respectivamente. A partir disso obtiveram a entropia relativa generalizada substituindo as funções q -logaritmo $\ln_q(\cdot)$ e logaritmo natural $\ln(\cdot)$ pelo inverso da exponencial deformada φ .

Dada uma exponencial deformada $\varphi : \mathbb{R} \rightarrow [0, \infty)$, a divergência entre duas distribuições de probabilidade $p = (p_i)$ e $q = (q_i)$ é expressa por

$$D_\varphi(p||q) = \sum_{i=1}^n \frac{\varphi^{-1}(p_i) - \varphi^{-1}(q_i)}{(\varphi^{-1})'(p_i)}, \quad (5)$$

onde temos $D_\varphi(p||q) = 0$ se $p = q$.

Conforme a equação (5), para o uso da entropia relativa generalizada, é necessário especificar exponencial deformada φ . Neste trabalho os testes são realizados utilizando duas funções exponenciais.

A primeira função exponencial utilizada é dada por

$$\varphi(x) = \ln(\exp(x) + 1), \quad (6)$$

cujas função inversa é

$$\varphi^{-1}(x) = \ln(\exp(x) - 1) \quad (7)$$

e derivada da função inversa

$$(\varphi^{-1})'(x) = \frac{\exp(x)}{\exp(x) - 1}. \quad (8)$$

A segunda função exponencial utilizada é a função κ -exponencial [19], onde para cada $-1 < \kappa < 1$ é obtida por

$$\exp_\kappa(x) = \exp \left(\int_0^x \frac{dt}{\sqrt{1 + \kappa^2 t^2}} \right), \quad (9)$$

de forma que

$$\exp_\kappa(x) = \begin{cases} (\kappa x + \sqrt{1 + \kappa^2 x^2})^{\frac{1}{\kappa}}, & \text{se } \kappa \neq 0 \\ \exp x, & \text{se } \kappa = 0 \end{cases}. \quad (10)$$

A inversa da função κ -exponencial é dada por

$$\ln_\kappa(x) = \begin{cases} \frac{x^\kappa - x^{-\kappa}}{2\kappa}, & \text{se } \kappa \neq 0 \\ \ln x, & \text{se } \kappa = 0 \end{cases} \quad (11)$$

cujas derivada da inversa é igual a

$$(\ln_\kappa)'(x) = \begin{cases} \frac{x^\kappa + x^{-\kappa}}{2} \frac{1}{x}, & \text{se } \kappa \neq 0 \\ \frac{1}{x}, & \text{se } \kappa = 0 \end{cases} \quad (12)$$

III. MÉTODO PROPOSTO

Os métodos de clusterização de Bregman são normalmente divididos em dois tipos: *hard* e *soft*. Enquanto na clusterização *soft* cada ponto de dados tem uma probabilidade de pertencer a uma das partições, na *hard* os grupos são definidos de modos que cada ponto de dados pertença a exatamente uma das partições. Neste trabalho é implementado um método adaptado da versão *hard* da clusterização Bregman utilizando uma φ -divergência para realizar a segmentação de regiões em imagens.

A. Clusterização Hard de Bregman

A segmentação de imagens por agrupamento é uma maneira de separar grupos de objetos onde pixels com mesmas características são agrupados em uma mesma região. Ou seja, dada uma imagem como um conjunto de observação (x_1, x_2, \dots, x_n) , em que cada pixel é dado como uma variável aleatória $X = \{\mathbf{x}_i\}_{i=1}^n$, conforme uma medida de probabilidade positiva ν , a segmentação é feita dividindo o conjunto X em k partições disjuntas $\{X_h\}_{h=1}^k$, de maneira que cada variável seja atribuída a apenas uma partição.

Definido o número de partições, inicializa o conjunto $M = \{\mu_h\}_{h=1}^k$, que assume valores aleatórios. Posteriormente é realizado a etapa de atribuição, onde para $\mathbf{x}_i \in X$ é associado $\mathbf{x}_i \in X_h$ com $h = \operatorname{argmin} d_\varphi(\mathbf{x}_i, \mu_h)$. Logo após é processado a etapa de re-estimação, onde é escolhido novos representantes do conjunto M por meio da média dos objetos do cluster correspondente. As etapas de atribuição e re-estimação são repetidas até que o critério de convergência seja obtido. Assim sendo, a execução do algoritmo de clusterização de Bregman (Algoritmo 1) [4] é finalizada após alcançar um número finito de iterações ou quando a perda total não pode mais ser

diminuída nem pela etapa de atribuição e nem alterando as médias dos clusters existentes. A perda de informação de Bregman é dada por:

$$L_\phi(M) = I_\phi(X) - I_\phi(M) = E_\pi[I_\phi(X_h)], \quad (13)$$

onde $L_\phi(M)$ pode ser interpretado como a informação de Bregman intra cluster, $I_\phi(M)$ como informação de Bregman inter clusters e $I_\phi(X)$ como informação total de Bregman que pode ser obtida por

$$I_\phi(X) = E_\nu[d_\phi(X, \mu)] = \sum_{i=1}^n \nu_i d_\phi(x_i, \mu). \quad (14)$$

Algorithm 1: K-means com φ -Divergência

Input: Conjunto $X = \{\mathbf{x}_i\}_{i=1}^n$;
 medida de probabilidade de ν sobre X ;
 divergência de Bregman d_ϕ e
 o número de clusters k
Output: M , Agrupamento *hard* $\{X_h\}_{h=1}^k$ de X ,
 minimizador local de
 $L_\phi(M) = \sum_{h=1}^k \sum_{\mathbf{x}_i \in X_h} \nu_i d_\phi(\mathbf{x}_i, \mu_h)$,
 onde $M = \{\mu_h\}_{h=1}^k$ de X

Método:

Inicializa $\{\mu_h\}_{h=1}^k$, podendo ser aleatoriamente
repetir
 {Etapa de atribuição}
para $i = 1$ até n **fazer**
 $X_h \leftarrow X_h \cup \{\mathbf{x}_i\}$, onde $h = \operatorname{argmin} d_\phi(\mathbf{x}_i, \mu_h)$
fim do para
 {Etapa de re-estimação}
para $h = 1$ até k **fazer**
 $\pi_h \leftarrow \sum_{\mathbf{x}_i \in X_h} \nu_i$
 $\mu_h \leftarrow \frac{1}{\pi_h} \sum_{\mathbf{x}_i \in X_h} \nu_i \mathbf{x}_i$
fim do para
até convergir
 $M \leftarrow \{\mu_h\}_{h=1}^k$

A divergência de Bregman mais simples e utilizada é a distância euclidiana quadrática, $d_\phi(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$, utilizada no algoritmo clássico K-means para a medição de similaridade.

IV. IMPLEMENTAÇÃO E RESULTADOS

A execução do método proposto é feito por meio da implementação da estrutura de clusterização *hard* (Algoritmo 1), substituindo a divergência de Bregman d_ϕ pela divergência D_φ (5) na medição de similaridade na segmentação de imagens. Na Fig. 1 é possível observar o resultado da segmentação utilizando o K-means clássico e utilizando a φ -divergência em uma imagem colorida. Em ambos métodos a clusterização é feita para cada camada RGB separada, considerando o número de clusters igual a 2 e na exponencial 2 é escolhido o valor de $\kappa = 0,4$.

A fim de analisar a performance do método de clusterização proposto é feito uma comparação com a segmentação utilizando o limiar Otsu, o algoritmo K-means e a divergência de Kullback-Leibler. Para realizar esta comparação é utilizado

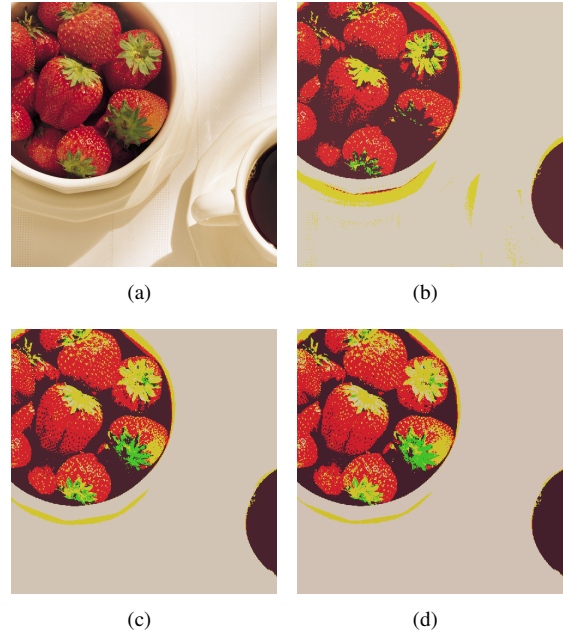


Fig. 1. Clusterização de imagem colorida [21]: (a) Imagem original; (b) K-means clássico; (c) φ (Exponencial 1); (d) φ (Exponencial 2).

o banco de dados NoisyOffice [20] que fornece imagens de textos escaneados com ruídos e seus *ground truth*. A principal ideia da segmentação de imagens de textos é dividir a imagem em região de texto e não-texto. Entretanto há imagens que apresentam ruídos que as intensidades de seus pixels estão mais próximas das intensidades dos textos do que das do fundo da imagem. Considerando isso, a clusterização é realizada dividindo a imagem em mais de duas regiões. Dessa forma é possível separar o que é texto, ruído e fundo, ou seja, 3 regiões.

Observa-se na Fig.2 as três regiões segmentadas separadamente pelo método proposto utilizando a exponencial 1 (6). Para cada região segmentada são atribuídos aos pixels pertencentes ao cluster o valor de seu centroide e aos outros pixels o valor 1, ou seja, para região não segmentadas é atribuído branco. Para a Fig.1 os centroides das regiões 1, 2 e 3 são de 0,0719, 0,4260 e 0,8455, respectivamente.

Na Fig. 3 observamos o comportamento da segmentação por limiarização e por clusterização para cada método e na Tab. I os seus resultados numéricos. O valor de κ utilizado é de 0,7. Vale salientar que este teste é realizado com a imagem sem utilizar nenhuma técnica de pré-processamento.

TABELA I

TAXA DE ACERTO DA SEGMENTAÇÃO POR CLUSTERIZAÇÃO

Método	Taxa de acerto
Limiar de Otsu	84.203%
K-means clássico	95.127%
Divergência KL	93.323%
Divergência φ (Exponencial 1)	93.777%
Divergência φ (Exponencial 2)	94.613%

Nota-se na Tab. I que o método K-means clássico, a divergência KL e o limiar de Otsu apresentam altos valores de taxa

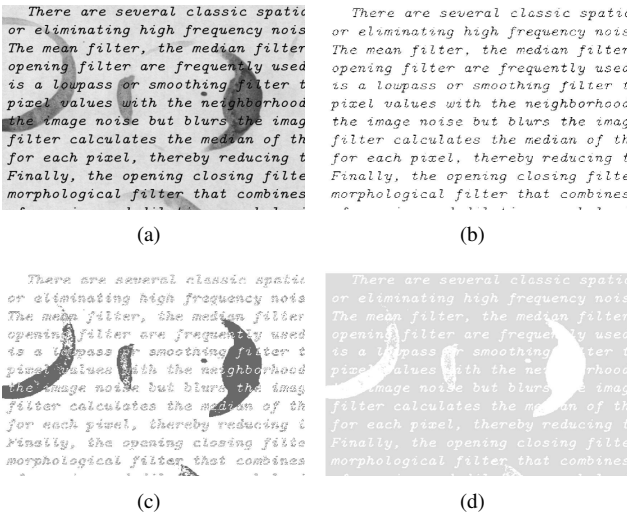


Fig. 2. Regiões segmentadas por clusterização: (a) Imagem original; (b) Região 1: Texto; (c) Região 2: Ruído; (d) Região 3: Fundo.

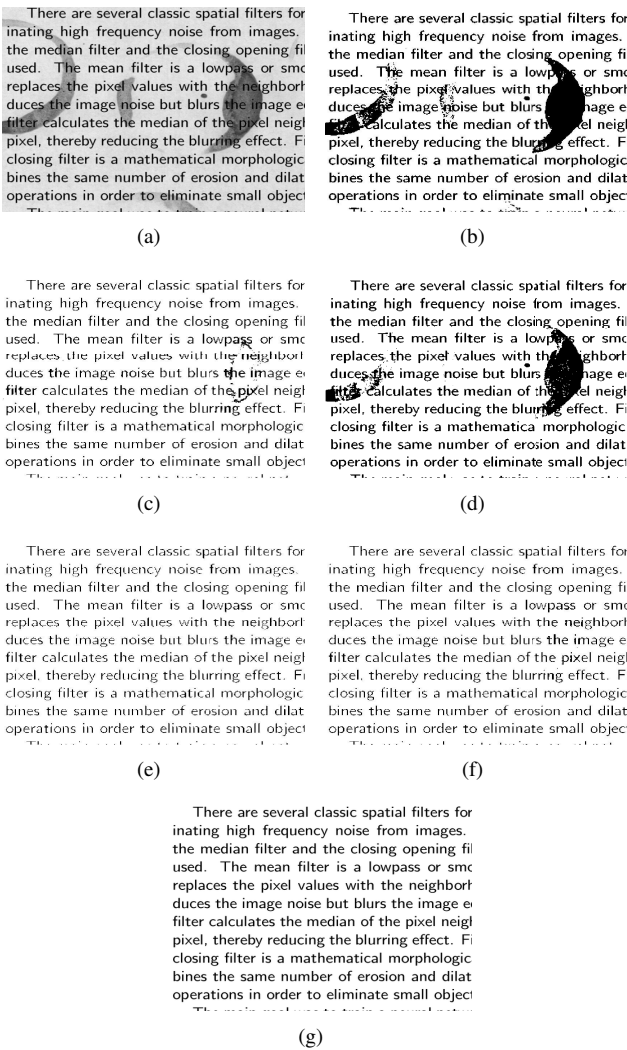


Fig. 3. Segmentação por clusterização: (a) Imagem original; (b) Limiar de Otsu; (c) K-means clássico; (d) Divergência KL (e) φ (Exponencial 1) (f) φ (Exponencial 2) (g) Ground truth.

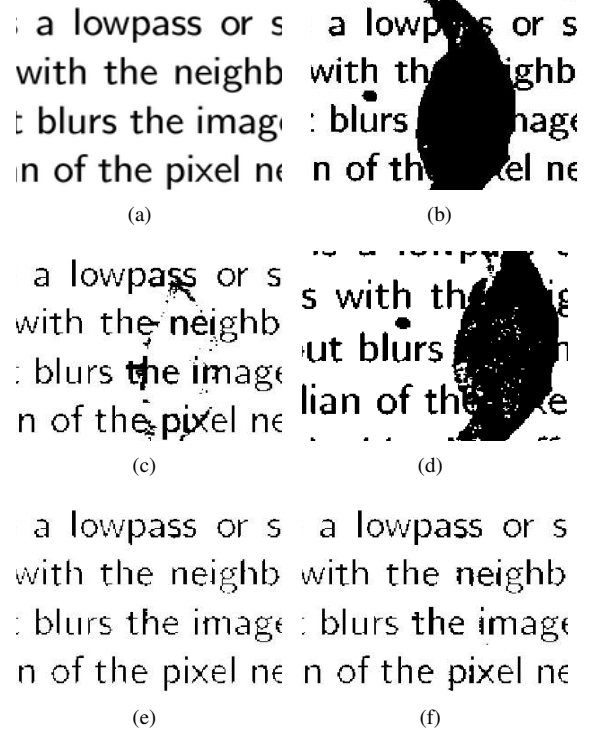


Fig. 4. Segmentação por clusterização: (a) Ground truth; (b) Limiar de Otsu; (c) K-means clássico; (d) Divergência KL (e) φ (Exponencial 1) (f) φ (Exponencial 2).

de acerto e pouca diferença em relação ao método proposto. Porém, analisando os resultados na Fig. 3 observa-se erros que podem comprometer o reconhecimento dos caracteres, visto que grande parte dos ruídos foram considerados como mesma região do texto. Comparado as duas exponenciais utilizadas com o *ground truth*, em ambas, parte do pixels do texto são agrupados como fundo da imagem, porém não o suficiente para comprometer a representação dos caracteres, sendo que dentre as duas a κ -exponencial se mostrou mais aplicável. É possível observar melhor a diferença na segmentação na Fig. 4, que apresenta um recorte da Fig. 3.

Considerando que $-1 < \kappa < 1$, na Tab. II é analisado a variação da taxa de acerto da segmentação para cada valor de κ na exponencial 2. É utilizado a mesma imagem da Fig. 3 com o mesmo número de clusters. A escolha do valor de κ é feita de forma empírica, considerando, além da taxa de acerto, o quão o texto é segmentado sem a presença de ruídos.

É analisado também o comportamento dos métodos de segmentação utilizados após passar por um pré-processamento onde é ajustado o contraste da imagem. Observa-se na Fig. 5 o resultado da segmentação e na Tab. III a taxa de acerto.

V. CONCLUSÕES

Este trabalho apresenta um novo método de calcular dissimilaridade na clusterização utilizando uma generalização das entropias relativas de Shannon e Tsallis, chamada de φ -divergência. Além disso, os métodos de clusterização baseados no limiar de Otsu, a divergência KL e o K-means clássico são comparados com o método proposto. Analisando suas

TABELA II
TAXA DE ACERTO DA CLUSTERIZAÇÃO DA DIVERGÊNCIA φ PARA CADA VALOR DE κ

κ	0	$\pm 0,1$	$\pm 0,2$	$\pm 0,3$	$\pm 0,4$	$\pm 0,5$	$\pm 0,6$	$\pm 0,7$	$\pm 0,8$	$\pm 0,9$
acerto(%)	94.40	95.26	95.84	95.728	95.32	93.93	92.19	94.61	94.72	92.90

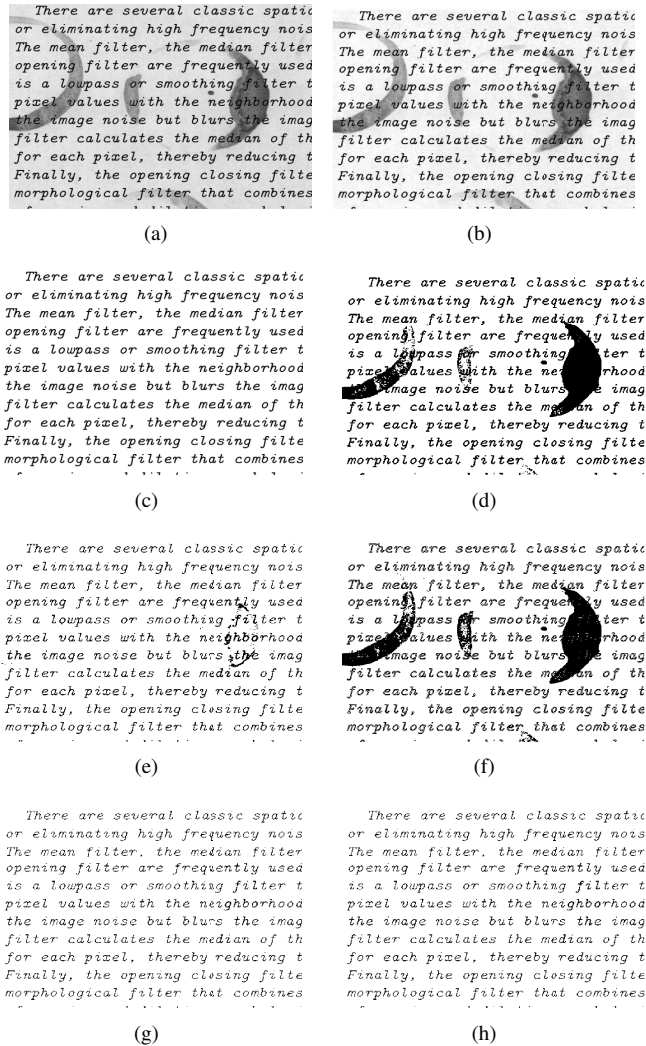


Fig. 5. Segmentação por clusterização de uma imagem pré-processada: (a) Imagem original; (b) Imagem processada; (c) Ground truth; (d) Limiar de Otsu; (e) K-means clássico; (f) Divergência KL (g) φ (Exponencial 1) (h) φ (Exponencial 2).

TABELA III
TAXA DE ACERTO DA SEGMENTAÇÃO POR CLUSTERIZAÇÃO DE UMA IMAGEM PRÉ-PROCESSADA

Método	Taxa de acerto
Limiar de Otsu	85.663%
K-means clássico	96.535%
Divergência KL	91.103%
Divergência φ (Exponencial 1)	95.116%
Divergência φ (Exponencial 2)	95.424%

performances é possível concluir que o método baseado na φ -divergência não só se mostrou aplicável como também pode apresentar melhor desempenho que métodos já conhecidos.

REFERÊNCIAS

- [1] I. A. Esquef, M. P. Albuquerque, e M. P. Albuquerque, *Processamento Digital de Imagens*, Centro Brasileiro de Pesquisas Físicas, 12 p. Rio de Janeiro, 2003.
- [2] R. Yogamangalam e B. Karthikeyan. *Segmentation Techniques Comparison in Image Processing*, International Journal of Engineering and Technology (IJET), v. 5, n. 1, pp. 307-313, 2013.
- [3] N. Otsu, *A Threshold Selection Method from Gray Level Histograms*, IEEE Transactions on Systems, Man, and Cybernetics, v.9, n.1, pp. 62-66, 1979.
- [4] A. B. S. Merugu, I. S. Dhillon e J. Ghosh, *Clustering with Bregman Divergences*, Journal of Machine Learning Research 6, pp. 1705-1749, 2005.
- [5] A. Buzo et al. *Speech coding based upon vector quantization*, IEEE Transactions on Acoustics, Speech and Signal Processing, v. 28, n. 5, pp. 562–574, 1980.
- [6] I. Dhillon, S. Mallela, e R. Kumar, *A divisive information-theoretic feature clustering algorithm for text classification*, Journal of Machine Learning Research, v. 3, n. 4, pp. 1265–1287, 2003.
- [7] C. E. Shannon, *A Mathematical Theory of Communication*, The Bell System Technical Journal, v. 27, pp. 379-423, 623-656, 1948.
- [8] C. Tsallis, *Possible Generalization of Boltzmann-Gibbs Statistics*, Journal of Statistical Physics, v. 52, n. 1-2, pp. 479-487, 1988.
- [9] S. Kullback e R. A. Leibler, *On Information and Sufficiency*, Ann. Math. Statist, v. 22, n. 1, pp. 79-86, 1951.
- [10] L. Borland, A. R. Plastino e C. Tsallis, *Information Gain within Nonextensive Thermostatistics*, Journal of Mathematical Physics, v. 39, n. 12, pp. 6490-6501, 1998.
- [11] R. F. Vigelis e C. C. Cavalcante, *On φ -Families of Probability Distributions*, Journal of Theoretical Probability, v. 26, n. 3, pp. 870–884, 2013.
- [12] R. F. Vigelis e C. C. Cavalcante, *Information Geometry: An Introduction to New Models for Signal Processing*, Signals and Images - CRC Press, pp. 455-491, 2015.
- [13] D. C. de Souza, C. C. Cavalcante e R. F. Vigelis, *Modified Maximum Likelihood Estimator*, 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), pp. 1-5, 2016.
- [14] J. A. Katkar e T. Baraskar, *Medical Image Segmentation using PCA and K-mean Clustering Algorithm*, iPGCON. Sangamner, 6 p, 2015.
- [15] F. A. Shmmla e W. Ashour, *Color Based Image Segmentation using Different Versions of K-Means in two Spaces*, Global Advanced Research Journal of Engineering, Technology and Innovation. v.1, n.9, pp. 30-41, 2013.
- [16] X. Zhao, Y. Li e Q. Zhao, *Mahalanobis Distance Based on Fuzzy Clustering Algorithm for Image Segmentation*, Digital Signal Processing, v.43, pp. 8-16, 2015.
- [17] H. Wei, L. Chen e L. Guo, *KL Divergence-Based Fuzzy Cluster Ensemble for Image Segmentation*, MDPI, entropy, v. 20, n.273, 17 p, 2018.
- [18] M. D. Gupta et al. *KL Divergence based Agglomerative Clustering for Automated Vitiligo Grading*, Computer Vision Foundation, pp. 2700 – 2709, Boston, 2015.
- [19] G. Pistone, *κ -exponential models from the geometrical viewpoint*, The European Physical Journal B, v.70, pp. 29–37, 2009.
- [20] F. Z. Martinez, S. E. Boquera e M. J. C. Bleda, *Behaviour-Based Clustering of Neural Networks Applied to Document Enhancement*, Computational and Ambient Intelligence. Springer, pp. 144-151, 2007.
- [21] R. C. Gonzalez e R. E. Woods, *Processamento Digital de Imagens*, 3. ed. [S.l.]: Prentice Hall, 624 p, 2011.