PROCEEDINGS OF SPIE

Hermite transform-based superpixel for texture image segmentation

Triana-Galeano, Vivian, González, Germán, Olveres, Jimena, Escalante-Ramírez, Boris

SPIE.

Hermite Transform-based Superpixel for Texture Image Segmentation

Vivian Triana-Galeano[a], Germán González[b], Jimena Olveres[b], and Boris Escalante-Ramírez[b]

[a]Posgrado en Ingeniería Eléctrica, Facultad de Ingeniería, Universidad Nacional Autónoma de México, Mexico City, Mexico
[b]Departamento de Procesamiento de Señales, Facultad de Ingeniería, Universidad Nacional Autónoma de México, Mexico City, Mexico

ABSTRACT

Superpixel algorithms oversegment an image by grouping pixels with similar local features such as spatial position, gray level intensity, color, and texture. Superpixels provide visually significant regions and avoid a large number of redundant information to reduce dimensionality and complexity for subsequent image processing tasks. However, superpixel algorithms decrease performance in images with high-frequency contrast variations in regions of uniform texture. Moreover, most state-of-the-art methods use only basic pixel information -spatial and color-, getting superpixels with low regularity, boundary smoothness and adherence. The proposed algorithm adds texture information to the common superpixel representation. This information is obtained with the Hermite Transform, which extracts local texture features in terms of Gaussian derivatives. A local iterative clustering with adaptive feature weights generates superpixels preserving boundary adherence, smoothness, regularity, and compactness. A feature adjustment stage is applied to improve algorithm performance. We tested our algorithm on Berkeley Segmentation Dataset and evaluated it with standard superpixel metrics. We also demonstrate the usefulness and adaptability of our proposal in medical image application.

Keywords: Hermite Transform, local clustering, medical image segmentation, Texture superpixels.

1. INTRODUCTION

The idea of superpixels was introduced by Ren and Malik[1] and describes the oversegmentation of an image into homogeneous regions that try to respect the image contours. These perceptually meaningful regions can be used to replace the structure of pixel grid[2] and have been used as a preprocessing in many computer vision applications such as contour detection,[3] segmentation,[4] object localization,[5] classification[6] and data augmentation in supervised training of deep neural networks.[7]

According to the literature, a superpixel method should include these properties: (i) Superpixels should be disjoint and assign a label to every pixel;[8] (ii) The clustering must group pixels into homogeneous areas in terms of its features;[9] (iii) Superpixels should adhere well to image boundaries;[2] (iv) Superpixels should be compact, placed regularly and exhibit smooth boundaries. The metrics to evaluate the superpixels properties include Undersegmentation Error (UE),[2,10] Achievable Segmentation Accuracy (ASA),[11] Explained Variation[8] and Global Regularity.[12]

Different approaches to generate superpixels have been proposed. We can classify these methods in different categories, according to Ref. 8, but the most used and important for this work are Contour Evolution[10] and Clustering-based methods such as SLIC superpixels which is the most representative superpixels method used in practical applications due to its time-efficiency.

However, most classic methods use only basic pixel information -spatial and color-, getting superpixels with low regularity, boundary smoothness and adherence, in regions where texture is present. For that reason, our proposed algorithm adds texture information to the common superpixel representation. These texture features

---

Further author information: (Send correspondence to V.T.G.)
V.T.G.: E-mail: viviantrianag@gmail.com

are obtained with the Hermite Transform, which has demonstrated to be useful for texture description[13, 14] especially in medical image applications.[15, 16]

In this paper, we use the Hermite Transform to include texture information in a SLIC superpixel method modification, to improve superpixel performance on images with important texture presence. Section 2 reviews the theoretical background of Hermite Transform. Section 3 describes the proposal feature representation and discriminability measure distance. In Section 4 Experiments and Results are presented. The final section concludes the paper.

2. HERMITE TRANSFORM

The Hermite Transform[17] is a local decomposition technique that expands an image into orthogonal polynomials with respect to a Gaussian window. The analysis functions are similar to Gaussian derivatives that, according to Ref. 18, fit the receptive field profiles of mammalian visual systems.

2.1 Polynomial Transform

A Polynomial Transform is a signal representation technique based on polynomial approximations within a local window. The input signal $L(x)$ is localized by multiplying it by a window function $V(x)$ and projecting it onto orthogonal polynomials with basis functions $G_n(x)$, where n is the polynomial grade. These basis functions are orthonormal with respect to $V^2(x)$.[17]

$$L_n(kT) = \int_{\infty}^{-\infty} L(x) \cdot G_n(x - kT)V^2(x - kT)dx \qquad (1)$$

Eq. 1 is the Direct Polynomial Transform which maps the input signal to the coefficients $L_n(kT)$ by convolving the input signal with the analysis functions or filter functions (eq. 2) following by a subsampling T.

$$D_n(x) = G_n(-x)V^2(-x) \qquad (2)$$

From the Inverse Polynomial Transform (eq. 3)[17] we can achieve signal reconstruction from the coefficients $L_n(kT)$, interpolating it with the Pattern functions (eq. 4) and summing over all orders n.

$$L(x) = \sum_{n=0}^{\infty} \sum_{k} L_n(kT) \cdot P_n(x - kT) \qquad (3)$$

$$P_n(x) = \frac{G_n(x)V(x)}{W(x)} \qquad (4)$$

2.2 Hermite Analysis Functions

The Hermite analysis functions of the Hermite transform[17] of degree n in one dimensional are defined in eq. 5:

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{\frac{-x^2}{\sigma^2}} \qquad (5)$$

where $H_n(x)$ are the Hermite polynomials given by Rodrigues' formula and σ is the standard deviation of the Gaussian window $V(x)$ (eq. 7). The filter function $Dn(x)$ is equal to the nth order derivative of a Gaussian (eq. 6).[17]

$$D_n(x) = \frac{1}{\sqrt{2^n n!}} \cdot \frac{d^n}{d(\frac{x}{\sigma})^n} \left[\frac{1}{\sigma\sqrt{\pi}} e^{\frac{-x^2}{\sigma^2}} \right] \qquad (6)$$

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \cdot e^{\frac{-x^2}{2\sigma^2}} \qquad (7)$$

2.3 Bidimensional Cartesian Hermite Transform

A Gaussian window in two dimensions has the property of being spatially separable[17] and the two dimensional analysis functions can be written as in equation 8.

$$D_{n-m,m}(x,y) = D_{n-m}(x)D_m(y) \tag{8}$$

where $n-m$ is the analysis order in direction x and m in direction y.

And input image $L(x,y)$ can be expanded into the basis $D_{n-m,m}(x,y)$ as shown in equation 9.

$$L_{n-m,m(x_0,y_0)} = \int_x \int_y L(x,y) \cdot D_{n-m,m}(x_0-x, y_0-y)dxdy \tag{9}$$

for $n = 0, 1, ..., d_{max}$ and $m = 0, ..., n$, where d_{max} is the maximum desired derivative degree.

Figure 1 shows the Coefficients of the Cartesian Hermite Transform where the vertical and horizontal patterns can be observed.
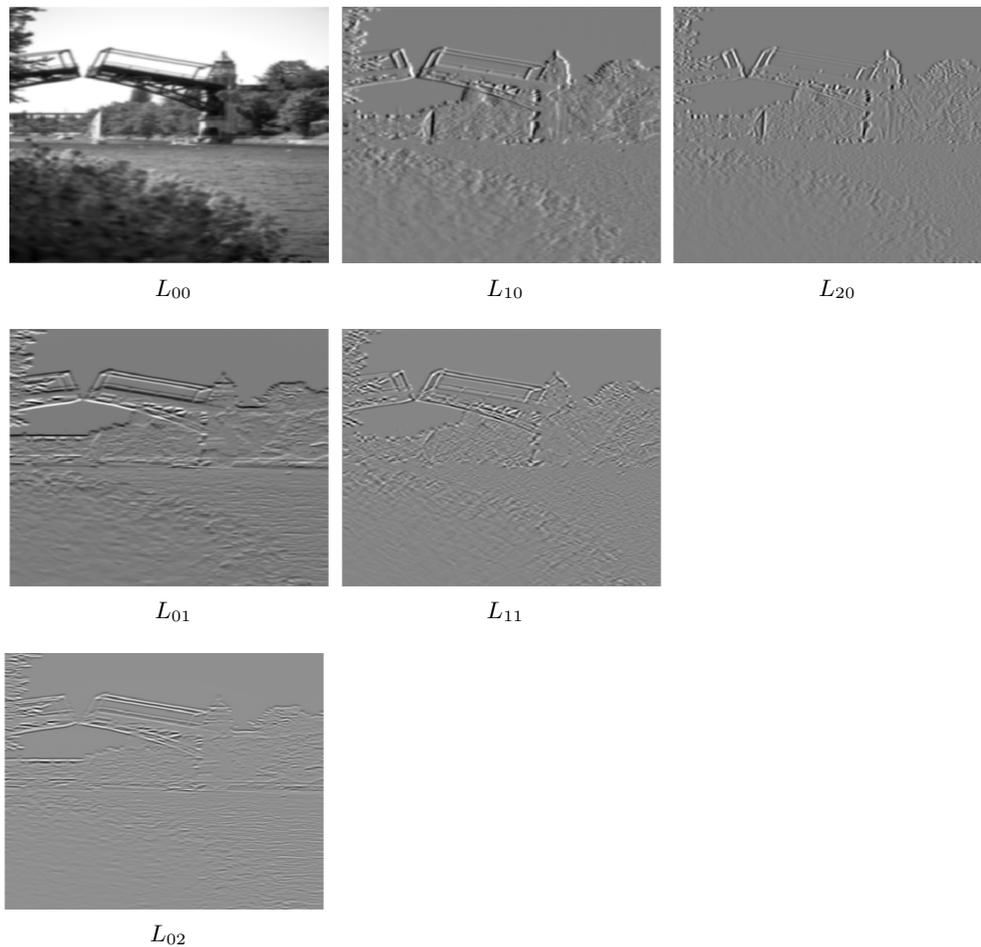

L_{00} L_{10} L_{20}

L_{01} L_{11}

L_{02}

Figure 1: Coefficients of the Cartesian HT up to second order for a image of Berkeley Dataset.

3. PROPOSAL

Most of the superpixel segmentation algorithms have important dependence on color and spatial features to get compact and perceptual segments. However, texture information on images is a great guideline for image segmentation, especially in medical image analysis.[15] For that reason, our proposal combines the idea of SLIC superpixels[2] with texture feature representation based on Hermite Transform,[13] including a weight distance adjustment to improve results.

3.1 Feature Representation

CIELAB is the color space used to calculate the color difference in many superpixel algorithms[2, 19] because it allows to perceive chromatic aberration easily. The color feature vector is $[l, a, b]$, where l stands for lightness, a and b are the color dimensions. The corresponding color difference d_c between two pixels p_i and p_j is calculated using a normalized Euclidean distance as follows:

$$d_c(p_i, p_j) = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \tag{10}$$

To enforce compactness, spatial feature difference is included (eq. 11). The spatial feature vector $[x, y]$ represents the vertical and horizontal coordinates of a pixel.

$$d_s(p_i, p_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{11}$$

In addition to the color and spatial features, texture feature represented by the Hermite Transform coefficients is included to improve the performance of the algorithm, especially in areas where color difference is not significant. The texture feature is represented by vector $tex = [L_{00}, L_{10}, L_{01}, L_{20}, L_{11}, L_{02}]$. The texture difference between p_i and p_j is set to:

$$d_{tex}(p_i, p_j) = \|tex_{p_i} - tex_{p_j}\| \tag{12}$$

Combining the color, spatial and texture feature, a pixel is described by:

$$p_i = [l_i, a_i, b_i, x_i, y_i, L_{00_i}, L_{10_i}, L_{01_i}, L_{20_i}, L_{11_i}, L_{02_i}] \tag{13}$$

3.2 Distance measure

The final distance measure between two pixels p_i and p_j described by eq. 13 is:

$$D = \sqrt{w_c(d_c)^2 + w_s(d_s)^2 + w_{tex}(d_{tex})^2} \tag{14}$$

where w_c, w_s and w_{tex} are weights to be adjusted iteratively. It is important to note that w_{tex} has the same length as vector tex. It means, each Hermite coefficient has its own weight.

As proposed in Ref. 20, the weights are adjusted using the sum of the within cluster distances SW_q for each feature (eq. 15).

$$SW_q = \sum_{r=1}^{k} \sum_{s=1}^{n} \widehat{u}_{p_s, c_r} d_q(p_q, c_r) \tag{15}$$

where q corresponds to the set of features, k is the number of superpixels, n is the number of pixels, $\widehat{u}_{p_s, c_r}$ is a binary variable to indicate whether a pixel p_s $(s = 1, 2, ...n)$ belongs to a cluster center $c_r(r = 1, 2, ...k)$; and $d_q(p_q, c_r)$ measures the distance of pixel p_s to center c_r on feature q.[20] Weight feature q is adjusted following eq. 16.

$$w_q = \frac{1}{\sum_{t \in FeatureSet} \left[\frac{SW_q}{SW_t} \right]^{\frac{1}{\beta - 1}}} \tag{16}$$

where β is fixed to 9, as specified by Ref. 21.

4. EXPERIMENTS AND RESULTS

We tested our algorithm on Berkely segmentation dataset[22] and compared our proposal with SLIC,[2] the most popular and used superpixel method, and Turbopixels,[10] one of the most regular and compact superpixel results.

4.1 Evaluation metrics

We evaluate the performance of the algorithm using standard superpixels metrics,[2, 8, 10, 12, 20, 23] such as Undersegmentation Error -UE, Achievable Segmentation Accuracy - ASA, Explained Variation - EV, and Global Regularity.[12]

For an image I, a superpixel decomposition $S = \{S_k\}_{k \in \{1,...,|S|\}}$ with $|S|$ superpixels S_k, and a ground truth denoted $G = \{G_j\}_{j \in \{1,...,|G|\}}$ with G_j a segmented region, we have the following evaluation metrics.

4.1.1 Undersegmentation Error - UE

The Undersegmentation Error - UE measures the overlap of superpixels with multiple and nearby ground truth segments.[8] Ref. 24 proposed a free parameter formulation of UE (equation 17).

$$UE(S,G) = \frac{1}{|I|} \sum_{S_k} \sum_{G_j} min\{|S_k \cap G_j|, |S_k \setminus G_j|\} \tag{17}$$

4.1.2 Achievable Segmentation Accuracy - ASA

The Achievable Segmentation Accuracy - ASA also aims at evaluating the overlap of superpixels with a ground truth. It is calculated as:

$$ASA(S,G) = \frac{1}{|I|} \sum_{S_k} \max_{G_j} |S_k \cap G_j| \tag{18}$$

4.1.3 Explained Variation -EV

Explained Variation - EV[8] helps to evaluate the homogeneity of the color clustering and is defined as equation 19.

$$EV(S) = \frac{\sum_{S_k} |S_k| [\mu(S_k) - \mu(I)]^2}{\sum_{p \in I} [I(p) - \mu(I)]^2} \tag{19}$$

4.1.4 Global Regularity - GR

The Global Regularity measure, proposed by Ref. 12, aims to evaluate the shape regularity and consistency in only one metrics.

$$GR(S) = SRC(S)SMF(S) \tag{20}$$

with,

Smooth Matching Factor - SMF compares the spatial distributions of the average superpixel shape S^* to each registered superpixel shape S_k^*.

$$SMF(S) = 1 - \sum_{S_k} \frac{|S_k|}{|I|} \cdot \left\| \frac{S^*}{|S^*|} - \frac{S^*}{|S_k^*|} \right\|_1 / 2 \tag{21}$$

Shape Regularity Criteria - SRC mesures the convexity, the smoothness of the contours and the balanced repartition of the pixels within the shape.

$$SRC(S) = \sum_{S_k} \frac{|S_k|}{|I|} CR(S_k) V_{xy}(S_k) \tag{22}$$

where, $CR(S) = \frac{CC(H_s)}{CC(S)}$ is the Criteria of Regularity - CR, with $CC(S) = |P(S)| \, |S|$ the relation between the perimeter and the area of a shape S. H_s is the convex hull which entirely contains the superpixel form S.[12] $V_{xy}(S) = \frac{min(\sigma_x, \sigma_y)}{max(\sigma_x, \sigma_y)}$ defines the variance as a ratio between the minimun and maximun variance of pixel positions x and y, which belong to S, where σ_x and σ_y are the standard deviations of the pixel positions.

4.2 Results

The quantitative evaluation of the behavior of the two methods -SLIC and Turbopixels- considered to compared our proposal with includes the metrics described in Section 4.1. We performed the experiments taking into account a number of superpixels variation from 100 to 900 with step of 100 superpixels. Figure 2 shows the results. According to these results, our proposal presents good homogeneity of color clustering Figure 2(c), acceptable respect of image objects Figure 2(a)(b) and low regularity Figure 2(d).
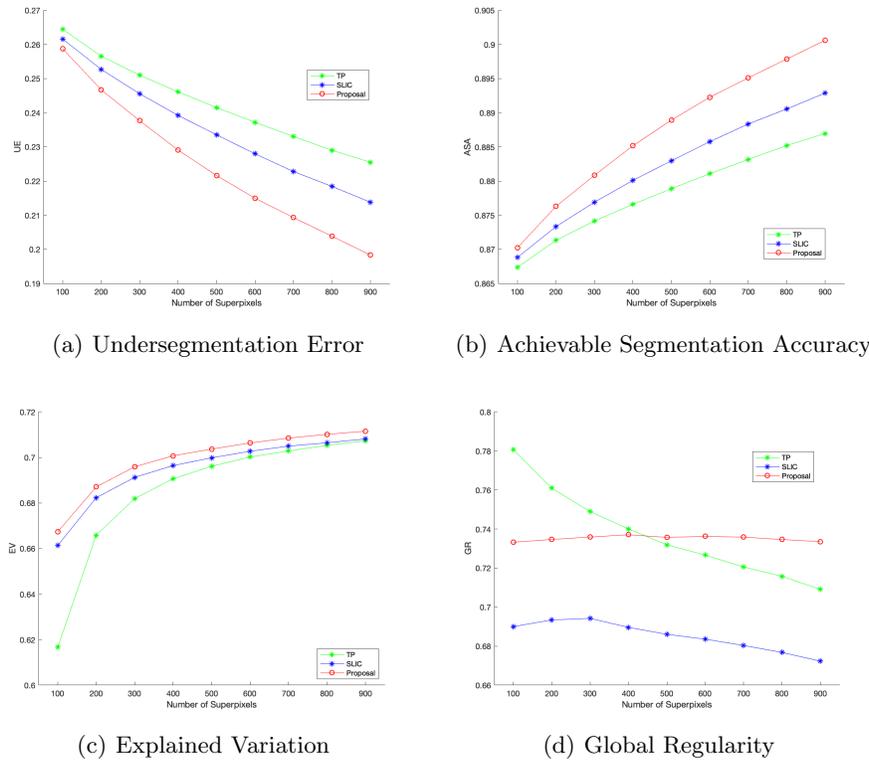

(a) Undersegmentation Error (b) Achievable Segmentation Accuracy

(c) Explained Variation (d) Global Regularity

Figure 2: Quantitative comparision for Berkeley Dataset images.

Figure 3 provides the visual comparison of superpixels generated with our proposal, SLIC and Turbopixel when the number of superpixels is 100. It can be seen that our proposal generates smoother superpixels than SLIC ones, but both algorithms present low regularity. Turbopixel generates the smoothest superpixels but its boundary adherence is low.
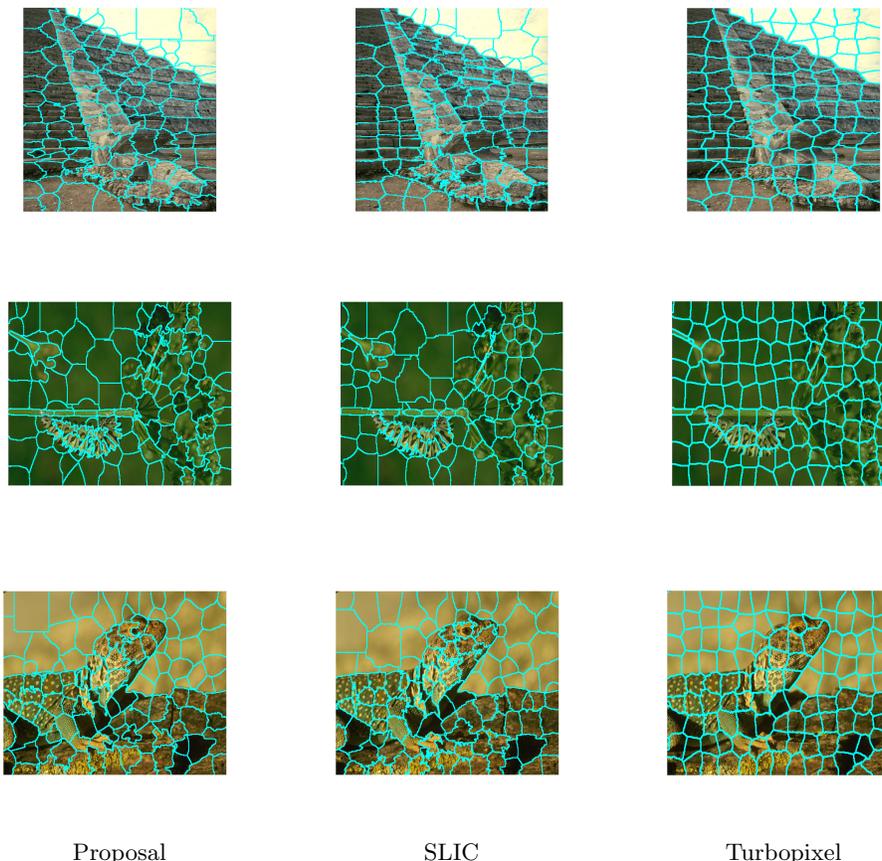

| Proposal | SLIC | Turbopixel |

Figure 3: Visual comparision for Berkeley Dataset images.

4.3 Application to medical imaging

Due to mammography limitations for breast cancer diagnosis, Ultrasound imaging has become an important aid to improve this diagnosis.[25] Shape and texture features are used to differentiate benign and malignant breast tumors. Because of that, better segmentation results of breast Ultrasound images have been sought. Superpixels have shown to be useful as a preprocessing step in image segmentation, and the Hermite Transform can provide good texture descriptors. Consequently, we propose to use our algorithm to presegment breast Ultrasound images to improve tumor classification.

The visual comparison of superpixels methods SLIC, Turbopixels and our proposal are shown in Figure 4 for breast Ultrasound images. The red contour is the expert delineation that is superposed to the superpixels contours (cyan contours) to confirm the usefulness of our superpixel proposal as a preprocessing step to improve the performance in subsequent medical image processing.

5. CONCLUSION

In this paper, we presented a superpixel algorithm that adds texture information, obtained with the Hermite Transform to the SLIC method with weights feature adjustment in each iteration, to improve the discrimination

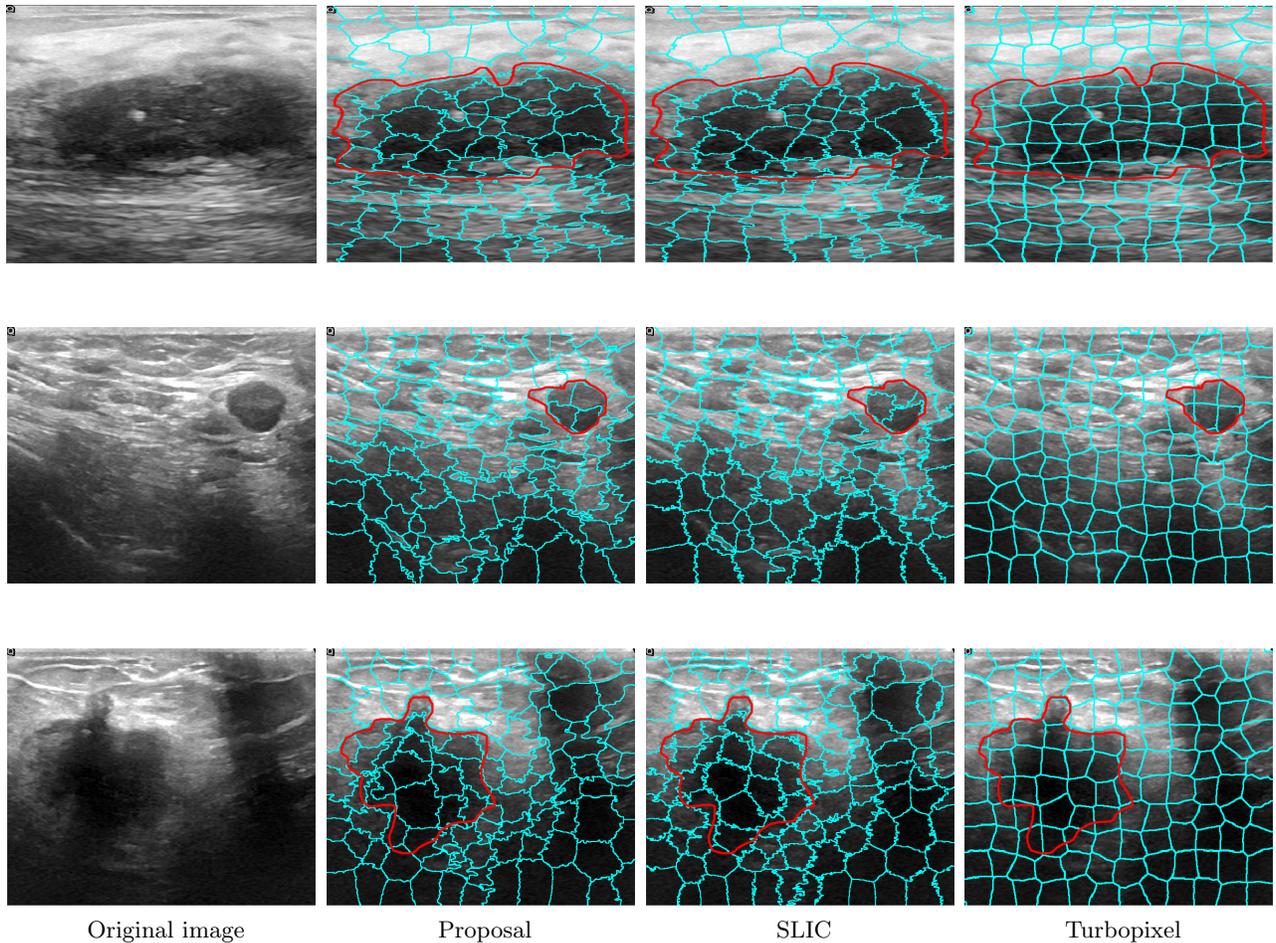| Original image | Proposal | SLIC | Turbopixel |

Figure 4: Visual result for Breast Ultrasound images.

of pixels in images of important texture content. Experimental results showed that we got superpixels with better object adherence, smoothness and homogeneity. Regularity remains a concern to improve, however, the superpixels are perceptually meaningful. The usefulness and adaptability of our proposal was demonstrated by applying it to breast Ultrasound images.

Texture feature inclusion with Hermite Transform and feature weights adjustment proved to be good but not enough to get better results in superpixels regularity. To cope with this issue, future work considers the idea of feature selection over the texture features in local regions of the images.

REFERENCES

[1] Ren, X. and Malik, J., "Learning a classification model for segmentation," in [*null*], 10, IEEE (2003).

[2] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S., "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence* **34**(11), 2274–2282 (2012).

[3] Arbelaez, P., Maire, M., Fowlkes, C., and Malik, J., "Contour detection and hierarchical image segmentation," *IEEE transactions on pattern analysis and machine intelligence* **33**(5), 898–916 (2010).

[4] Borovec, J., Švihlík, J., Kybic, J., and Habart, D., "Supervised and unsupervised segmentation using superpixels, model estimation, and graph cut," *Journal of Electronic Imaging* **26**(6), 061610 (2017).

[5] Fulkerson, B., Vedaldi, A., and Soatto, S., "Class segmentation and object localization with superpixel neighborhoods," in [*2009 IEEE 12th international conference on computer vision*], 670–677, IEEE (2009).

[6] Maghsoudi, O. H., "Superpixel based segmentation and classification of polyps in wireless capsule endoscopy," in [*2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*], 1–4, IEEE (2017).

[7] Zhang, Y., Yang, L., Zheng, H., Liang, P., Mangold, C., Loreto, R. G., Hughes, D. P., and Chen, D. Z., "Spda: superpixel-based data augmentation for biomedical image segmentation," *arXiv preprint arXiv:1903.00035* (2019).

[8] Stutz, D., Hermans, A., and Leibe, B., "Superpixels: An evaluation of the state-of-the-art," *Computer Vision and Image Understanding* **166**, 1–27 (2018).

[9] Giraud, R., Ta, V.-T., Papadakis, N., and Berthoumieu, Y., "Texture-aware superpixel segmentation," in [*2019 IEEE International Conference on Image Processing (ICIP)*], 1465–1469, IEEE (2019).

[10] Levinshtein, A., Stere, A., Kutulakos, K. N., Fleet, D. J., Dickinson, S. J., and Siddiqi, K., "Turbopixels: Fast superpixels using geometric flows," *IEEE transactions on pattern analysis and machine intelligence* **31**(12), 2290–2297 (2009).

[11] Liu, M.-Y., Tuzel, O., Ramalingam, S., and Chellappa, R., "Entropy rate superpixel segmentation," in [*CVPR 2011*], 2097–2104, IEEE (2011).

[12] Giraud, R., Ta, V.-T., and Papadakis, N., "Evaluation framework of superpixel methods with a global regularity measure," *Journal of Electronic Imaging* **26**(6), 061603 (2017).

[13] Estudillo-Romero, A., Escalante-Ramirez, B., and Savage-Carmona, J., "Texture analysis based on the hermite transform for image classification and segmentation," in [*Optics, Photonics, and Digital Technologies for Multimedia Applications II*], **8436**, 843619, International Society for Optics and Photonics (2012).

[14] Estudillo-Romero, A. and Escalante-Ramirez, B., "Rotation-invariant texture features from the steered hermite transform," *Pattern Recognition Letters* **32**(16), 2150–2162 (2011).

[15] Vargas-Quintero, L., Escalante-Ramírez, B., Marín, L. C., Huerta, M. G., Cosio, F. A., and Olivares, H. B., "Left ventricle segmentation in fetal echocardiography using a multi-texture active appearance model based on the steered hermite transform," *computer methods and programs in biomedicine* **137**, 231–245 (2016).

[16] Olveres, J., Nava, R., Escalante-Ramírez, B., Vallejo, E., and Kybic, J., "Left ventricle hermite-based segmentation," *Computers in biology and medicine* **87**, 236–249 (2017).

[17] Martens, J.-B., "The hermite transform-theory," *IEEE Transactions on Acoustics, Speech, and Signal Processing* **38**(9), 1595–1606 (1990).

[18] Young, R. A., "Orthogonal basis functions for form vision derived from eigenvector analysis," *Investigative Ophthalmology & Visual Science (ARVO Suppl.)* **19**, 172 (1978).

[19] Li, Z. and Chen, J., "Superpixel segmentation using linear spectral clustering," in [*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*], 1356–1363 (2015).

[20] Xiao, X., Zhou, Y., and Gong, Y.-J., "Content-adaptive superpixel segmentation," *IEEE Transactions on Image Processing* **27**(6), 2883–2896 (2018).

[21] Huang, J. Z., Ng, M. K., Rong, H., and Li, Z., "Automated variable weighting in k-means type clustering," *IEEE transactions on pattern analysis and machine intelligence* **27**(5), 657–668 (2005).

[22] Martin, D., Fowlkes, C., Tal, D., and Malik, J., "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in [*Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*], **2**, 416–423, IEEE (2001).

[23] Machairas, V., Faessel, M., Cárdenas-Peña, D., Chabardes, T., Walter, T., and Decencière, E., "Waterpixels," *IEEE Transactions on Image Processing* **24**(11), 3707–3716 (2015).

[24] Neubert, P. and Protzel, P., "Superpixel benchmark and comparison," in [*Proc. Forum Bildverarbeitung*], **6**, 1–12 (2012).

[25] Daoud, M. I., Atallah, A. A., Awwad, F., Al-Najjar, M., and Alazrai, R., "Automatic superpixel-based segmentation method for breast ultrasound images," *Expert Systems with Applications* **121**, 78–96 (2019).