

LOCAL MULTIPLE ORIENTATIONS ESTIMATION USING K-MEDOIDS

Zhanghui Kuang, Guodong Pan, Kwan-Yee K. Wong

Department of Computer Science
University of Hong Kong
Pokfulam Road, Hong Kong

ABSTRACT

Estimation of local multiple orientations plays an important role in many image processing and computer vision tasks. It has been shown that the detection of orientations in an image patch corresponds to fitting multiple axes to its Fourier transform. In this paper, k-medoids are introduced to detect local multiple orientations in the Fourier domain. Medoids are related to a well-known matrix eigenvector problem. A hierarchical schema with eigensystem and energy distribution analysis is employed to determine the number of orientations in an image patch. The proposed approach detects two types of orientation structure (ridges and edges) without difference. Experimental results on synthetic and real images show that the proposed method can detect multiple orientations with high accuracy and is robust against noise.

Index Terms— Image processing, local multiple orientations, k-medoids

1. INTRODUCTION

Image local orientation is an important feature used in many image processing and computer vision tasks, such as texture analysis, edge detection, and image segmentation [1]. Local single orientation can be used to describe image features like lines and edges [2]. Local multiple orientations (corners, T-junctions, Y-forks, and X-sings) are widely used in applications like image registration [3], motion estimation [4] and tracking [5], and they can describe textures and fabrics in a more explicit manner in comparison to single one. This paper presents a unified framework of multiple orientations detector to overcome the restriction of single orientation detectors.

Many techniques for orientation estimation have been proposed in the literature. Most local orientation estimation approaches are based on the analysis of the spatial domain. Rao and Schunck [6] divided a texture image into an angle image and a coherence image using an oriented filter, namely the gradient of Gaussian, and performed manipulations on resulting gradient vector field. How to deal with noise is the major problem that gradient-based methods have to face because local gradients are very sensitive to noise, making the estimation of local orientations rather unreliable [7]. Freeman

and Adelson [8] introduced steerable filters to calculate the filter response at various orientations. Similarly, Paris et al. [9] employed Canny, Gabor, and Gaussian second derivative filters with different orientations to detect the local orientations of hair images. Nevertheless, they are often limited in precision [7].

In parallel with research effort focused in the spatial domain, Knutsson and Granlund [1] showed that the frequency properties of an image patch reflect changes in its gradient direction in the spatial domain, and they introduced quadrature filters to detect local orientations. Bigun et al. [10] posed the orientation estimation problem as a least squares fitting of an axis in the Fourier domain. However, these two approaches only deal with single orientation and errors arise when handling complex texture.

In order to deal with image features occluded at object boundaries which go beyond single local orientation, Aach et al. [11] introduced a framework of double orientations estimation. However, modeling all the signals as the superposition of two oriented signals is still too restrictive. For instance, plant roots generate "Y-forks" and this implies that three independent orientations have to be modeled. Muhllich and Aach [2, 11] developed a mathematical model for multiple orientations, which allows a unified treatment of additively and occludingly superimposed structures as well as combinations of these. Nevertheless, their method is restricted to superimposed model. Vliet and Faas [12] employed a clustering method to divide gradient vectors into at most 3 clusters. Since one cluster collects the background pixels and each cluster was analyzed separately using the structure tensor, their method only handles two orientations.

Two types of orientations, namely ridges (a light-dark-light signal orthogonal to the orientation direction or vice versa) and edges (only a single light-dark transition or vice versa), can be treated in the Fourier domain without difference since frequency reflects the changes in intensity of an image independent of the spatial location. Noise can be relatively easily removed using a Gaussian envelope. Different from the approaches discussed above, this paper introduces k-medoids to detect local multiple orientations in the frequency domain. Radial vectors are weighted by the corresponding energy of Fourier transform, and then k-medoids are utilized

to fit axes to the weighted vectors of the Fourier transform, where the medoid of each cluster is related to the solution of a well-known matrix eigenvector problem. Eigensystem and energy distribution analysis are employed to determine the number of orientations. Experimental results show that the proposed method can obtain multiple orientations with high accuracy and is robust against noise. Application such as estimation of hair orientation field is discussed.

The rest of the paper is organized as follows: Section 2 models and analyzes the local multiple orientations problem. Main steps of the proposed algorithm is given in Section 3. Performance evaluation of the proposed algorithm is described in Section 4. Section 5 is devoted to conclusions.

2. MODELING LOCAL MULTIPLE ORIENTATIONS PROBLEM

The frequency properties of an image patch reflect the changes in intensities. If an image patch has a preference for some orientation, most of the energy in the Fourier domain will be concentrated in a sector oriented at the same angle as the gradients inside the patch. The more orientation specific the patch is, the narrower the sector will be [1]. Therefore, to detect local multiple orientations in the spatial domain is consequently the same as to check for the existence of energy concentration along lines (highlight lines) in the Fourier domain. Two types of orientations including ridges and lines can be treated without difference as frequency reflects changes in intensity of image independent of the spatial location.

Let $f(x, y)$, with $f : R^2 \rightarrow R$, denote a gray-level image and w be a $m \times n$ (m and n are odd and set to be 11 in our experiments) patch with its center at (x_0, y_0) . Transforming w into the Fourier domain and shifting direct component (DC) to the center gives $F(u, v)$ ($u = -(m-1)/2, \dots, 0, \dots, (m-1)/2, v = -(n-1)/2, \dots, 0, \dots, (n-1)/2$). In practice, both w and $F(u, v)$ are weighted by a Gaussian envelope with variance $0.03n$ to emphasize the central pixel and eliminate noise, respectively. Since this weighting does not influence our considerations, we drop it for ease of notation. Each term of the Fourier transform can be further represented by a vector $\omega_{uv} = \|F(u, v)\|_2 \begin{pmatrix} u \\ v \end{pmatrix}$. More weight is assigned to a term with high energy and far away from the origin $(0, 0)$. (see Fig. 1)

Vector ω can also be represented as (l, θ) in polar coordinates, where l and θ are amplitude and angle of the vector, respectively. Sorting ω according to its angle gives the ordered set $\{\omega_i | i = 1, \dots, N\}$, where $N = mn$. To detect local K orientations is the same as finding K axes to fit its Fourier transform with a minimal error. These axes can be given by K unit vectors $\mu_j = (1, \bar{\theta}_j)$. Briefly, it minimizes

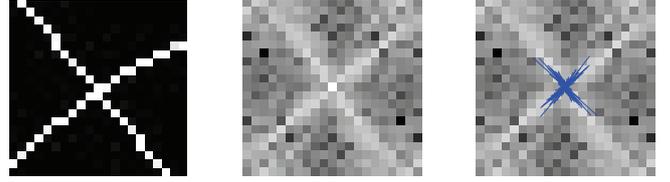


Fig. 1. Example of an image patch and its corresponding Fourier transform and vector representation. Left: original image patch. Middle: Fourier transform of the left. Right: vector representation of the middle.

the following distortion measure

$$E = \sum_{i=1}^N \min_{1 \leq j \leq K} (\nu(\omega_i, \mu_j)) \quad (1)$$

where $\nu(\omega_i, \mu_j)$ is the dissimilarity between two vectors ω_i and μ_j . It can be defined as following

$$\nu(\omega_i, \mu_j) = -l_i^\tau \cos^2(\theta_i - \bar{\theta}_j) \quad (2)$$

where τ denotes the impact of the amplitude of vector ω_i . Usually, τ is set to be 2, and the dissimilarity term would become the negative value of the square of the projection from vector ω_i onto the unit axis μ_j .

The objective function E is a piecewise trigonometric function with multiple local minima, resulting in a difficult optimization.

3. K-MEDOIDS FOR LOCAL MULTIPLE ORIENTATIONS

K-medoids algorithm, being the most common generalization of k-mean, is widely used for its simplicity [13]. It aims at partitioning observations into clusters in which each observation belongs to the cluster with the nearest medoid. Minimizing the objective function (1) is the same as minimizing

$$E = \sum_{i=1}^N \sum_{j=1}^K r_{ij} \nu(\omega_i, \mu_j) \quad (3)$$

where r_{ij} are indicators ($r_{ij} \in \{0, 1\}$) describing which of the K clusters the vector ω_i is assigned to.

An iteration approach is applied to minimize E . Consider first the determination of the r_{ij} with fixed μ_j . Since E in (3) is a linear function of r_{ij} and the terms involving different i are independent, it can be optimized for each i separately by choosing r_{ij} to be 1 for whichever value of j that gives the minimum value of $\nu(\omega_i, \mu_j)$. More formally, this can be expressed as

$$r_{ij} = \begin{cases} 1 & \text{if } j = \arg \min_k \nu(\omega_i, \mu_k) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

After each vector ω_i is assigned to a cluster, the medoids should be updated, i.e., optimization of the μ_j with the r_{ij} held fixed. Since μ_j is independent, it can be computed separately by minimizing the following

$$e(\mu_j) = \sum_{i=1}^N r_{ij} \nu(\omega_i, \mu_j) \quad (5)$$

Let $q_j = (\cos(\bar{\theta}_j), \sin(\bar{\theta}_j))^T$. Eq. (5) is equivalent to $-q_j^T C_j q_j$, where

$$C_j = \begin{bmatrix} \sum_{i=1}^N r_{ij} l_i^T \cos^2(\theta_i) & \frac{1}{2} \sum_{i=1}^N r_{ij} l_i^T \sin(2\theta_i) \\ \frac{1}{2} \sum_{i=1}^N r_{ij} l_i^T \sin(2\theta_i) & \sum_{i=1}^N r_{ij} l_i^T \sin^2(\theta_i) \end{bmatrix} \quad (6)$$

Minimizing $-q_j^T C_j q_j$ is a standard eigenvalue problem, which is discussed in [7]. The optimal vector q_j is the unit eigenvector of C_j corresponding to the largest eigenvalue.

The process of building r_{ij} and updating μ_j is repeated until the medoid of each cluster converges. Finally, the local multiple orientations of pixel (x_0, y_0) in an image can be achieved at $\{\bar{\theta}_1 + \frac{\pi}{2}, \dots, \bar{\theta}_k + \frac{\pi}{2}\}$.

Since corners, T-junctions, Y-forks, X-sings are the most frequent features of interests in many practical situations and unstable and little useful results will arise when too many orientations are being considered, the maximal number of local orientations estimated is set to be 3. A hierarchical schema is introduced to test whether or not to accept the zero-, single-, double- or triple-orientation hypotheses. Clearly, high ratio between $F(0, 0)$ and all the energy of Fourier transform indicates a homogeneous neighborhood without a clear oriented structure. Confidence for the zero-orientation assumption is high if

$$\frac{\|F(0, 0)\|_2}{\sum_u \sum_v \|F(u, v)\|_2} > \varepsilon_1 \quad (7)$$

where ε_1 is a confidence parameter.

In regions where the zero-orientation model is rejected ($K \geq 1$), we test other hypotheses using eigensystem analysis. The single-orientation is first tested where the number of orientation K is set to be 1. Two eigenvalues λ_1 and λ_2 ($\lambda_1 \geq \lambda_2$) of C_1 are obtained. Since eigenvalues are positively correlated with the energy in the corresponding axis directions, they can be used as a measure of confidence of the hypothesis. The energy independent quantity $R = (\lambda_1 - \lambda_2)/(\lambda_1 + \lambda_2)$ will be suitable for this task [10]. Confidence for the single-orientation assumption is high if

$$R > \varepsilon_2 \quad (8)$$

where ε_2 is a threshold for judging whether the number of orientations is more than one.

Table 1. Mean errors for local multiple orientations estimation in the noise-free image and images corrupted by noise

Noise	Noise-free	Gaussian	Gaussian	Gaussian	Gaussian
SNR	$+\infty$ dB	6.5dB	5.5dB	4.9dB	4.6dB
Error	0.05rad	0.05rad	0.07rad	0.08rad	0.09rad
Noise	Poisson	Uniform	Uniform	Uniform	Uniform
SNR	41.5dB	5.7dB	5.4dB	5.0dB	4.7dB
Error	0.05rad	0.06rad	0.07rad	0.07rad	0.08rad

The case that (8) does not hold indicates more than one orientations exist in the neighborhood ($K \geq 2$). Continue to test triple-orientation assumption ($K = 3$), getting $\lambda_1^1, \lambda_2^1, \lambda_1^2, \lambda_2^2, \lambda_1^3, \lambda_2^3$, where λ_i^j are the eigenvalues of C_j and $\lambda_1^j \geq \lambda_2^j$. Obviously, $\lambda_1^j + \lambda_2^j = \text{Trace}(C_j)$, which is positively correlated with the energy that is concentrated in the j^{th} orientation. If the assumed number of orientations is too many, then energy distribution does not appear uniformly. Therefore, a suitable confidence criterion for the triple-orientation assumption is

$$\frac{\min_{1 \leq j \leq 3} (\lambda_1^j + \lambda_2^j)}{\max_{1 \leq j \leq 3} (\lambda_1^j + \lambda_2^j)} \geq \varepsilon_3 \quad (9)$$

If (9) does not hold, then double-orientation assumption is confidence.

4. EXPERIMENTS AND APPLICATION

The proposed approach is evaluated using both synthetic and real images. Experimental results show that the proposed method can detect multiple orientations with high accuracy and is robust against noise. Applications such as estimation of orientation field as well as feature points of hair are discussed.

Using a specially constructed test image, the technique proposed previously is evaluated in terms of accuracy and robustness against noise. Fig. 2 shows the synthetic image, which consists of sine waves with decreasing amplitude in the radial direction, including all possible directions. Gaussian uncorrelated white noise, Poisson noise, and uniform noise are added into noise-free images, respectively. The orientations of points (including the ones with single-, double, or triple-orientations) in each image are estimated. The number of local orientations can be estimated correctly and local multiple orientations can be obtained with little errors even the image is polluted by noise. The mean errors for orientations in images corrupted by different kinds of noise with various levels are given in Table 1. Fig. 2 shows the estimation of local multiple orientations of a noise-free image and an image with Gaussian noise at SNR 4.6 dB, respectively.

Fig. 3 shows the performance of the proposed approach and the algorithm proposed by Aach et. al. [11]. Our method can judge the number of orientations of all the points detected

correctly while Aach's cannot. The parameter setting is: $\varepsilon_1 = 0.9$, $\varepsilon_2 = 0.5$ and $\varepsilon_3 = 0.88$.

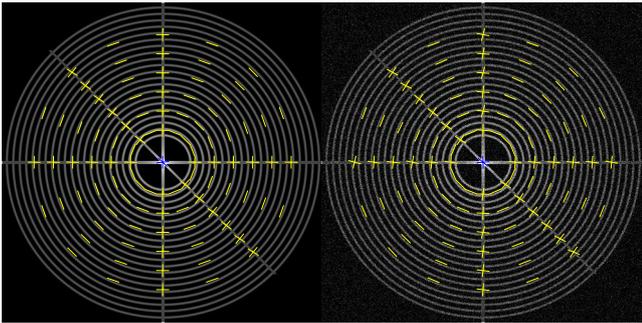


Fig. 2. Synthetic image for local multiple orientations detection. Left: orientations of a noise-free image. Right: orientations of an image with Gaussian noise at SNR 4.6 dB.

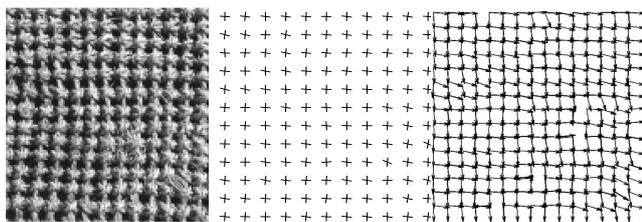


Fig. 3. Multiple orientations estimation applied to texture image. Left: texture image. Middle: orientations of the left obtained by the proposed approach. Right: orientations of the left obtained by [11].

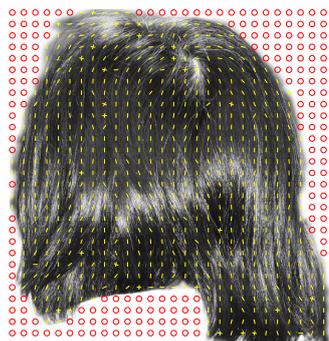


Fig. 4. Estimate the orientation field of hair

To show the robust performance and wide usefulness of the proposed algorithm, it is utilized to estimate the orientations of hair which is a challenging problem [14] since hair fibers are smaller than a pixel, and aliasing and hair lighting properties make it difficult to predict any strong properties [9]. Fig. 4 shows an image of hair. The proposed algorithm can divide the hair image into background (zero-orientation, marked by \circ), smooth partition of hair (single-orientation,

marked by line segment in its corresponding orientation) and mess partition of hair (more than two orientations, marked by \times).

5. CONCLUSION

K-medoids are introduced to detect local multiple orientations of images in the Fourier domain. Medoids are related to the solutions of a well-known matrix eigenvector problem. A hierarchical schema is employed to estimate the number of orientations using eigensystem and energy distribution analysis. The proposed approach can obtain multiple orientations with high accuracy and is robust against noise. Experiment results on texture image show that it performs better in estimation of the number of orientations compared with other existing methods. An application in capturing the orientation field of hair is also illustrated.

6. REFERENCES

- [1] H. Knutsson and G. H. Granlund, "Texture analysis using two-dimensional quadrature filters," in *IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, 1983, pp. 206–213.
- [2] M. Muhlich and T. Aach, "Analysis of multiple orientations," *IEEE Transactions on Image Processing*, vol. 18, no. 7, pp. 1424–1437, 2009.
- [3] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [4] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.
- [5] N. C. Stache, T. H. Stehle, M. Mhlich, and T. Aach, "Towards multiple-orientation based tensor invariants for object tracking," in *European Signal Processing Conference*, Firenze, 2006, pp. 1–14.
- [6] A. R. Rao and B. G. Schunck, "Computing oriented texture fields," *Graphical Models and Image Processing*, vol. 53, no. 2, pp. 157–185, 1991.
- [7] X. G. Feng, "Analysis and approaches to image local orientation estimation," M.S. thesis, 2003.
- [8] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.
- [9] S. Paris, H. M. Briceno, and F. X. Sillion, "Capture of hair geometry from multiple images," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 712–719, 2004.
- [10] J. Bigun, G. H. Granlund, and J. Wiklund, "Multidimensional orientation estimation with applications to texture analysis and optical-flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 775–790, 1991.
- [11] T. Aach, C. Mota, I. Stuke, M. Muhlich, and E. Barth, "Analysis of superimposed oriented patterns," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3690–3700, 2006.
- [12] L. J. van Vliet and F. G. A. Faas, "Multi-orientation analysis by decomposing the structure tensor and clustering," in *Proceedings of the 18th International Conference on Pattern Recognition*, 2006, pp. 856–860.
- [13] C. M. Bishop, *Pattern Recognition and machine learning*, Springer, 2007.
- [14] Y. C. Wei, E. Ofek, L. Quan, and H. Y. Shum, "Modeling hair from multiple views," *ACM Transactions on Graphics*, vol. 24, no. 3, pp. 816–820, 2005.