LAPLACE GRADIENT BASED DISCRIMINATIVE AND CONTRAST INVERTIBLE DESCRIPTOR

Zhenwei Miao[⋆†], *Kim-Hui Yap*[⋆], *Xudong Jiang*[⋆], *Subbhuraam Sinduja*[⋆], *Zhenhua Wang*[†]

[⋆†]IIM Ltd, Singapore 138628,
miaozhenwei@gmail.com
[⋆]School of Electrical and Electronic Engineering,
Nanyang Technological University, Singapore 639798,
{ekhyap, exdjiang, sinduja001}@ntu.edu.sg
[†]Rachel and Selim Benin School of Computer Science and Engineering,
Hebrew University of Jerusalem, Israel 91905,
zhwang.me@gmail.com

ABSTRACT

The performance of local descriptors such as SIFT drops under severe illumination changes. In this paper, we propose a Discriminative and Contrast Invertible (DCI) local feature descriptor. In order to increase the discriminative ability of the descriptor under illumination changes, a Laplace gradient based histogram is proposed. Moreover, a robust contrast flipping estimate is proposed based on the divergence of a local region. Experiments on fine-grained object recognition and retrieval applications demonstrate the superior performance of the DCI descriptor to others.

***Index Terms*—** SIFT descriptor, contrast-invertible, image retrieval, logo search, face recognition

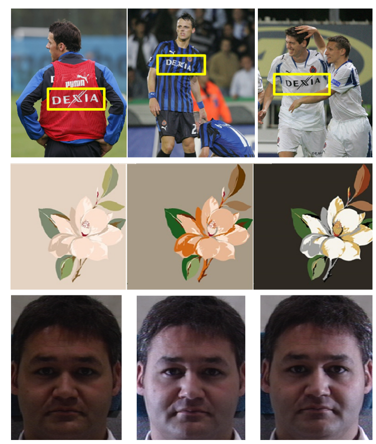


Fig. 1. Sample images under different illumination conditions. The first row shows the logos with contrast inversion, the second row shows the wallpaper with different colors and third rows shows the same face with different lighting.

1. INTRODUCTION

Local feature description [1–6] has drawn intensive attention in image processing, analysis and search. In general, local patches extracted by various interest point detectors [6–13] are expected to be described by descriptors that are discriminative and tolerant to geometrical and illuminative variations. Despite efforts have been devoted to develop kinds of filters and descriptors [14–18], the performances of existing descriptors are still limited especially under conditions where illumination changes greatly. Such changes occur frequently in real-world scenarios as shown in Fig. 1. In such cases, previous descriptors encounter difficulties in recognition and retrieving the images with the identical objects yet different illumination conditions.

Generally, the contrast changes of images can be classified into two types: bright-dark order preserved changes and bright-dark order disturbed changes. In images with bright-dark order preserved changes, the relative order of the pixel intensity remains the same while the image patch becomes either brighter or darker following linear and/or nonlinear transformations. In contrast, images with bright-dark disturbed changes will not retain the relative order. The majority of previous local feature descriptors, such as SIFT [6], GLOH [5] and PCA-SIFT [19], are designed for the first type. But they fail to address the issues caused by severe nonlinear illumination changes and bright-dark order disturbed changes [18]. The mirror and inversion invariant SIFT (MI-SIFT) [20], MAX-SIFT [21], ExHoG [22] and DRLBP [23] were proposed to tackle the bright-dark order disturbed changes. However, they cannot solve the problem of partial illumination changes.

In this work, we propose a novel descriptor named Discriminative and Contrast Invertible (DCI) descriptor. The DCI is designed to enhance the performance of local feature descriptors under illumination changes and contrast inver-

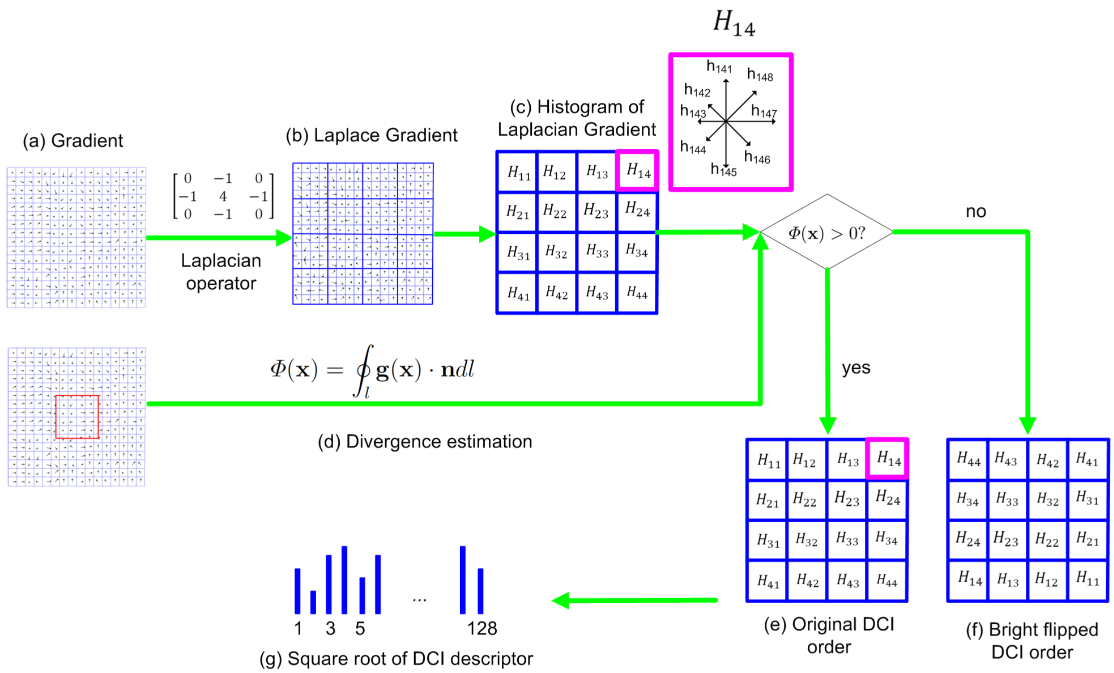


Fig. 2. Flowchart of the proposed DCI descriptor. (a) gradient of image patch, (b) Laplace gradient, (c) histogram of Laplace gradient in the 4×4 blocks, (d) divergence-based brightness flipping estimator, (e) original DCI order, (f) DCI order under brigh-dark flip case, and (g) DCI descriptor after L_1 normalization and square rooting.

sions. We propose the Laplace gradient to alleviate the lighting variation problem. The DCI descriptor is formed by concatenating the histograms of the Laplace gradient in different bins. Moreover, to address the bright/dark inversion problem, we propose an inversion-estimation function based on the divergence of gradient. The proposed descriptor is evaluated across diverse object visual search tasks, including searching logos, commercial design images and faces. Experimental results demonstrate that the proposed DCI descriptor outperforms the state-of-the-art descriptors.

2. THE PROPOSED DESCRIPTOR

Fig. 2 illustrates the work flow of the DCI descriptor. As how the SIFT descriptor is developed, a local patch with the standard size of 31×31[1] is extracted around each interest point at the given scale and aligned along its dominant orientation [6]. Following that, the gradient of the image patch is extracted as shown in Fig. 2(a). The Laplace gradient shown in Fig. 2(b) is computed to enhance the discriminative ability and robustness to brightness changes. Then, the Laplace gradient map in 2(b) is divided into 4×4 sub-regions. The distribution of the Laplace gradient in each sub-region is quantized into an eight-orientation-bin histogram to represent each sub-region in Fig. 2(c). To handle the bright-dark order disturbed changes, the convergence based contrast flipping estimator in Fig. 2(d) is generated. According to the sign of the estimator, either the original order version in Fig. 2(e) or the inverted order in Fig. 2(f) is chosen to encode the descriptor. Lastly, the histograms in each block are concatenated into one vector. Rooting algorithm is applied to further enhance the descriptor. The rooted histogram as shown in Fig. 2(g) after L_1 normalization forms the final DCI descriptor.

[1]Here we follow the default of the SIFT descriptor provided in vlfeat http://www.vlfeat.org/

2.1. Laplace Gradient

To tackle the problem caused by illumination changes and to enhance the discriminative ability of the gradient based descriptor, we propose to use higher order surface information such as the curvature rather than the lower order information. Considering that isotropic operators are helpful in capturing the more invariant information, we propose to employing Laplace gradient to describe the local patch. Let the image gradient $\mathbf{g}(\mathbf{x})$ at location $\mathbf{x}$ be

$$\mathbf{g}(\mathbf{x}) = \nabla I(\mathbf{x}) = \frac{\partial I(\mathbf{x})}{\partial x}\overrightarrow{i} + \frac{\partial I(\mathbf{x})}{\partial y}\overrightarrow{j}, \quad (1)$$

where $\mathbf{x} = \{x, y\}$ is the coordinate, $I(\mathbf{x})$ is the pixel intensity and $\overrightarrow{i}$ and $\overrightarrow{j}$ are the directions along x and y axises. By applying the scalar Laplacian operator to the two components of the gradient image independently, the Laplace gradient is arrived at

$$\begin{aligned}\mathbf{d}(\mathbf{x}) &= \nabla^2 \mathbf{g}(\mathbf{x}) \quad (2)\\ &= \Big(\frac{\partial^3 I(\mathbf{x})}{\partial x^3} + \frac{\partial^3 I(\mathbf{x})}{\partial x \partial y^2}\Big)\overrightarrow{i} + \Big(\frac{\partial^3 I(\mathbf{x})}{\partial y \partial x^2} + \frac{\partial^3 I(\mathbf{x})}{\partial y^3}\Big)\overrightarrow{j}. \quad (3)\end{aligned}$$

The Laplacian operator can boost the discriminative information of local features by the high order derivative.

2.2. Divergence based Contrast Flipping

To address the bright-dark flipping issue, we propose to use the divergence of the local region to check whether the local region is bright or dark. A bright blob is a region that the gradient converges to the center whereas a dark blob is the region that the gradient diverges from the center. The divergence of gradient is defined as

$$\mathrm{div}\mathbf{g}(\mathbf{x}) = \nabla \mathbf{g}(\mathbf{x}) = \nabla^2 I(\mathbf{x}). \tag{4}$$

The equation (4) indicates that the divergence of the gradient is the Laplacian of the image. Thus, it is plausible that the blob can be detected by the response of the Laplacian filter. The response of the Laplacian filter produces the largest signal to noise ratio at the center of the blob region as its shape matches with that of the blob. Therefore, it is conceivable that the divergence is more reliable in detecting whether this region is bright or dark. The surface integration of the divergence is

$$\varPhi(\mathbf{x}) = \iint_s \mathrm{div}\mathbf{g}(\mathbf{x}) ds. \tag{5}$$

The implementation of the surface integration is clear. It can be easily computed from the gradient. Based on the Stokes' theorem [24], the integration of the divergence gradient over a surface equals to the integration of the gradient over the boundary of the local region, that is

$$\varPhi(\mathbf{x}) = \iint_s \mathrm{div}\mathbf{g}(\mathbf{x}) ds = \oint_l \mathbf{g}(\mathbf{x}) \cdot \mathbf{n} dl, \tag{6}$$

where $\mathbf{n}$ is the outward-pointing unit normal vector on the boundary.

2.3. The DCI Descriptor

With the Laplace gradient, an additional step to tackle the illumination changes is adopting the rooting algorithm [25]. The DCI descriptor is derived as follows. Let the concatenation of the Laplace gradient in all the 4×4 sub-regions be $H = \{h_1, ..., h_{128}\}$. The L_1 normalized H is $H_n = \{h_{1n}, ..., h_{128n}\}$ where $h_{in} = h_i / \sum H$. The DCI descriptor H_{nr} is the square root of the elements in H, defined as

$$H_{nr} = \{\sqrt{h_{1n}}, ..., \sqrt{h_{128n}}\}. \tag{7}$$

3. EXPERIMENTS

In this section, we evaluate our proposed DCI descriptor with comparison to the state-of-the-art SIFT [6], LIOP [18], MI-SIFT (MIS) [20], MAX-SIFT (MAS) [21] and RIDE [1] descriptors in the applications of logo visual search, wallpaper visual search, and face recognition.

Table 1. Retrieval Performance on the BelgaLogos Database.

Logos	SIFT	LIOP	MIS	MAS	RIDE	DCI
	100%	100%	87.3%	92.9%	100%	**100%**
	22.0%	4.3%	27.9%	25.8%	39.8%	**40.4%**
	61.9%	37.4%	60.7%	66.6%	68.5%	**68.8%**
	24.6%	5.7%	8.1%	10.8%	24.3%	**31.6%**
	68.8%	49.3%	82.5%	85.6%	87.0%	**87.7%**
	67.8%	51.1%	61.0%	66.3%	67.5%	66.5%
mAP	57.5%	41.3%	54.6%	57.8%	64.5%	**65.8%**

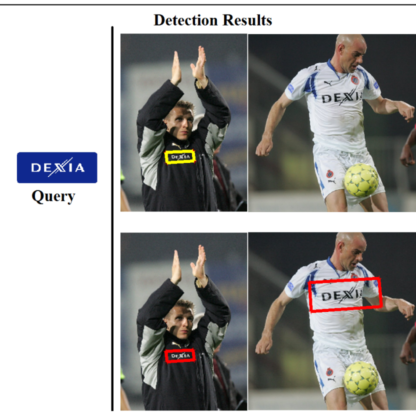


Fig. 3. Logo detection results based on the SIFT and DCI descriptors. The first row shows the results obtained using SIFT marked with yellow bounding box, and the second row illustrates the results obtained using DCI marked with red bounding box. The proposed DCI can effectively detect both two logos in these two images but SIFT cannot.

3.1. Logo Visual Search

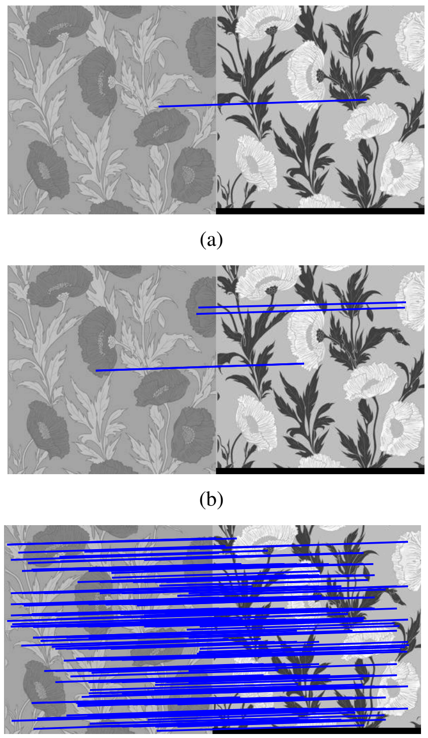
The descriptors are evaluated on the challenging BelgaLogos database [26] that contains 10,000 real-world images from various events as shown in the first row of Fig. 1. Hessian-Laplacian detector [27] with the default setting from vl-feat[2] is used to extract interest points from each image. The brute-force search algorithm is used for all the descriptors evaluated here. The evaluation is tested on the widely evaluated six logos [28], the 'US-President', 'Mercedes', 'Kia', 'Ferrari', 'Dexia' and 'Base'.

The Average Precision (AP) of each logo and the mean Average Precision (mAP) for all six logos are given in Table 1. It shows that the DCI descriptor outperforms other descriptors most of the cases. It increases the mAP by more than 8% compared to the SIFT descriptor. In order to explain the performance improvement of the DCI descriptor, visual inspections on the logo detection are given in Fig. 3. It shows that the SIFT descriptor can detect the logo in the top left image but the top right image where the bright and dark parts are the inversion of query. As expected, our proposed DCI descriptor can effectively detect the logo from both images under such severe contrast changes.

[2] The vl-feat is downloaded from http://www.vlfeat.org/

Table 2. Retrieval Performance on the Wallpaper Database.

	SIFT	LIOP	MIS	MAS	RIDE	DCI
mAP (%)	64.0	49.5	61.5	67.6	64.5	**74.0**



(a)

(b)

(c)

Fig. 4. Correct matched pairs obtained by using the (a) SIFT, (b) LIOP, and (c) DCI descriptors. The blue lines represent the correctly matched pairs between the query and reference images.

3.2. Wallpaper Visual Search

A wallpaper dataset as shown in the second row of Fig. 1 is chosen to evaluate the proposed descriptor as this dataset includes a considerable number of images with various illumination variations from the same design. In total, the reference image set contains 522 images from 77 categories that are provided by a wallpaper design company. The test image set contains 1,014 images. The wallpaper search algorithm in [29] is implemented here. The image retrieval results in Table 2 show that the proposed DCI improves the retrieval performance by 6% compared with the other descriptors. It demonstrates that the DCI is more robust to the severe contrast changes which widely exist in the wallpaper dataset.

As shown in Fig. 4, eyeballing two images with large illumination changes can find that both the SIFT and LIOP descriptors identify only limited number of correctly matched pairs in the event of severe illumination changes. In contrast, the proposed DCI descriptor can identify much more correctly matched pairs.

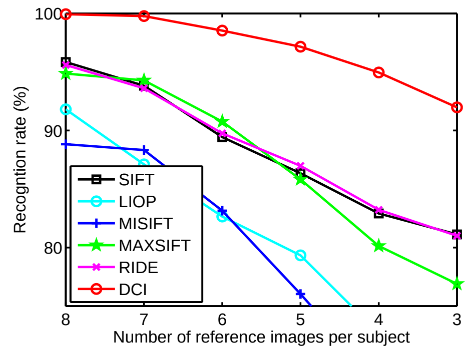


Fig. 5. Face recogntion rate versus number of reference images on the Multi-PIE database.

3.3. Face Recognition

Over years, face recognition is always an active research topic [30–33]. In this experiment, the proposed DCI descriptor is evaluated on one of the widely accepted face databases: the CMU Multi-PIE database [34]. The CMU Multi-PIE database contains face images captured in 4 sessions with variations in illumination, expression and pose. For the purpose of this experiment, the face image sets with illumination variation is selected.

The first 105 subjects that appear in all 4 sessions are used. Images are cropped and down-sampled to 100×82 pixels. In total, 20 neutral-expression images with different illumination are used for the evaluation, which produce $20 \times 4 \times 105 = 8400$ images for the evaluation. For each subject, the first t images are selected as the reference and the rest $20 - t$ images are used as the query. The ℓ_0-LoG detector [8] which is an illumination invariant interest point detector is used to extract the keypoints from the face images. The algorithm of image recognition through interest point matching is adapted from [6]. Experimental results are shown in Fig. 5. It shows that the proposed DCI descriptor outperforms other descriptors over all cases. When the reference number of images is 3, the DCI descriptor achieves more than 10% incrementation compared to other descriptors.

4. CONCLUSIONS

We propose a local feature descriptor named DCI that is discriminative and contrast invertible in illumination changes and contrast inversion. The Laplace gradient is computed to describe each pixel. A divergence-based contrast flipping estimator is created for images with the bright-dark order disturbed variations. The square root of the histogram of Laplace gradient after L_1 normalization is used to further mitigate the problems caused by illumination changes. Experiments on the BelgaLogos, Wallpaper, CMU Multi-PIE databases exhibit the superior performance of the DCI descriptor in object search applications over the state-of-the-art descriptors. It suggests that the DCI descriptor is robust to the illumination changes and contrast inversion.

5. REFERENCES

[1] L. X. Xie, J. D. Wang, W. Y. Lin, B. Zhang, and Q. Tian, “Ride: Reversal invariant descriptor enhancement,” in *International Conference on Computer Vision*, 2015.

[2] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “Orb: an efficient alternative to sift or surf,” in *IEEE International Conference on Computer Vision*, 2011, pp. 2564–2571.

[3] W. L. Zhao and C. W. Ngo, “Flip-invariant sift for copy and object detection,” *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 980–991, 2013.

[4] T. Tuytelaars and K. Mikolajczyk, “Local invariant feature detectors: a survey,” *Fundations and Trends in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2008.

[5] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[6] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[7] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, “Brief: Binary robust independent elementary features,” *European Conference on Computer Vision*, pp. 778–792, 2010.

[8] Z. W. Miao, X. D. Jiang, and K.-H Yap, “Contrast invariant interest point detection by zero-norm log filter,” *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 331–342, 2016.

[9] R. Kimmel, C. P. Zhang, A. M. Bronstein, and M. M. Bronstein, “Are mser features really interesting?,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2316–2320, 2011.

[10] Z. W. Miao and X. D. Jiang, “Interest point detection using rank order LoG filter,” *Pattern Recognition*, vol. 46, no. 11, pp. 2890–2901, November 2013.

[11] Z. W. Miao and X. D. Jiang, “A novel rank order LoG filter for interest point detection,” in *International Conference on Acoustics, Speech and Signal Processing*, 2012, pp. 937–940.

[12] Z. W. Miao and X. D. Jiang, “A vote of confidence based interest point detector,” in *International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 1459–1463.

[13] E. Rosten, R. Porter, and T. Drummond, “Faster and better: A machine learning approach to corner detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, 2010.

[14] X. D. Jiang, “Iterative truncated arithmetic mean filter and its properties,” *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1537–1547, 2012.

[15] Z. W. Miao and X. D. Jiang, “Further properties and a fast realization of the iterative truncated arithmetic mean filter,” *IEEE Transactions on Circuits and Systems Part II: Express Briefs*, vol. 59, no. 11, pp. 810–814, November 2012.

[16] Z. W. Miao and X. D. Jiang, “Weighted iterative truncated mean filter,” *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4149–4160, August 2013.

[17] Z. W. Miao and X. D. Jiang, “Additive and exclusive noise suppression by iterative trimmed and truncated mean algorithm,” *Signal Processing*, vol. 99, pp. 147–158, June 2014.

[18] Z. H. Wang, B. Fan, and F. Wu, “Local intensity order pattern for feature description,” in *IEEE International Conference on Computer Vision*, 2011, pp. 603–610.

[19] Y. Ke and R. Sukthankar, “Pca-sift: A more distinctive representation for local image descriptors,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 506–513.

[20] R. Ma, J. Chen, and Z. Su, “Mi-sift: mirror and inversion invariant generalization for sift descriptor,” in *Proceedings of the ACM International Conference on Image and Video Retrieval*, 2010, pp. 228–235.

[21] L. X. Xie, Q. Tian, and B. Zhang, “Max-sift: Flipping invariant descriptors for web logo search,” in *IEEE International Conference on Image Processing*, 2014, pp. 5716–5720.

[22] A. Satpathy, X. D. Jiang, and H.-L. Eng, “Human detection by quadratic classification on subspace of extended histogram of gradients,” *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 287–297, 2014.

[23] A. Satpathy, X. D. Jiang, and H.-L. Eng, “Lbp-based edge-texture features for object recognition,” *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 1953–1964, 2014.

[24] George Springer, *Introduction to Riemann surfaces*, vol. 473, Addison-Wesley Reading, Mass., 1957.

[25] R. Arandjelovic and A. Zisserman, “Three things everyone should know to improve object retrieval,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2911–2918.

[26] A. Joly and O. Buisson, “Logo retrieval with a contrario visual query expansion,” in *ACM International Conference on Multimedia*, 2009, pp. 581–584.

[27] K. Mikolajczyk and C. Schmid, “Scale & affine invariant interest point detectors,” *International journal of computer vision*, vol. 60, no. 1, pp. 63–86, 2004.

[28] Y. Jiang, J. Meng, and J. Yuan, “Randomized visual phrases for object search,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3100–3107.

[29] K.-H Yap and Z. W. Miao, “Hybrid feature-based wallpaper visual search,” in *IEEE International Symposium on Circuits and Systems*, 2015, pp. 730–733.

[30] Z. W. Miao, W. Ji, Y. Xu, and J. Yang, “A novel ultrasonic sensing based human face recognition,” in *IEEE Ultrasonics Symposium*, 2008, pp. 1873–1876.

[31] Z. W. Miao, W. Ji, Y. Xu, and J. Yang, “Human face classification using ultrasonic sonar imaging,” *Japanese Journal of Applied Physics*, vol. 48, no. 7S, pp. 1–5, 2009.

[32] X. D. Jiang, B. Mandal, and A. Kot, “Eigenfeature regularization and extraction in face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 383–394, 2008.

[33] C. Geng and X. D. Jiang, “Fully automatic face recognition framework based on local and global features,” *Machine Vision and Applications*, vol. 24, no. 3, pp. 537–549, 2013.

[34] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, “Multi-pie,” *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.