# Rotation-Invariant Categorization of Colour Images using the Radon Transform

Andrew P. Papliński
Clayton School of Information Technology
Monash University, VIC 3800, Australia
Email: Andrew.Paplinski@monash.edu

*Abstract*—We have derived a novel rotation-invariant feature vector, or signature, for two-variable functions like images. The feature vector is calculated as an angular integral of the Radon transform of the function. Three such feature vectors are calculated for each colour image. Subsequently, these feature vectors are used to categorize colour images in a network of 3+1 self-organizing modules. The 3-D 'labels' produced by the first level modules are used by the second level modules and can be thought of as a "universal neuronal code". The network is trained for un-rotated images and then tested for rotated images. It has been demonstrated that rotation of images by the angles included in calculation of the Radon transform results in the perfect categorization. For the angles in between, that is, those not included in the Radon transform, a small shift in categorization might occur, keeping, however, the objects well inside their clusters. Since calculation of the rotation-invariant feature vectors is very simple and involves only summations of signals (pixel values), hence very fast, it is postulated that such a mechanism might be included in the biological vision systems.

*Index Terms*—Computer vision, Rotation invariance, Radon transform, Image categorization, Self-organizing networks, Signature function

## I. INTRODUCTION

This paper is continuation and generalization of considerations presented in [1] where we investigated the problem of rotational invariance of binary images of characters. Rotation and scale invariance, a special case of the affine transformation, is the central issue of object recognition in computer vision. The current most popular and dominating directions in invariant visual recognition are based on a variety of local descriptors (e.g. [3], [4], [5]). Three groups of methods, namely, the Scale Invariant Feature Transform [6] (SIFT), PCA-SIFT [7] and the "Speeded Up Robust Features" (SURF) [8], seem to dominate the field. A comparison between the three methods is given in [9]. In [10] a rotation-invariant kernels are discussed in application to shape analysis. All the above methods deliver impressive results in object recognition. Although the speed of the methods has improved significantly, by their nature, the algorithms used are multistep recursive procedures and as such, are difficult to be considered as good candidates for models of human vision, where the speed is a survival issue.

In our work we concentrate on the problem of categorization of objects. The key to object categorization is the problem of identifying the image features that are used to classify the objects into categories using some learning technique. At the sensory, or receptive field level, we use features based on the Radon transform that seem to work very well in the problem of rotation invariant categorization of object.

The Radon transform has a long history of application in computer tomography, and relatively recently has been applied in a variety of image processing problems. The first to our knowledge such an application is presented in [2] where the Radon transform and SVD is used to create features used by the classifier. Subsequently, the Radon transform is used in conjunction with other transforms, wavelet and Fourier included. Magli et al. [11] and Warrick and Delaney [12] seem to initiate the use of the Radon transform in combination with the wavelet transform. More recently, a similar combination of transforms has been used in rotation invariant texture analysis [13], [14], and in shape representation [15]. Other approach to rotation invariant texture analysis uses the Radon transform in combination with the Fourier transform [16]. Chen and Kégl [17] consider feature extraction using combination of three transforms: Radon, wavelet and Fourier. In [18], texture classification is performed by using a feature descriptor based on the Radon transform and an affine invariant transform. Miciak [19] describes a character recognition system based on the Radon transform and the Principal Component Analysis (PCA). Hejazi et al. [20] present the discrete Radon transform in rotation invariant image analysis. Close to our considerations are object identification problems discussed by Hjouj and Kammler in [21].

In the above papers the reader can find many variants of detailed description of the Radon transform and its properties. Here, we can only reiterate the basic fact that Radon transform, $Rf(\theta, s)$, is composed of sums of pixels along the line that crosses the visual object under the angle $\theta$ at the distance $s$ from the origin. It can be noted that the Radon transform of the image rotated by a known angle, $\alpha$, can be easily inferred from the transform of the un-rotated image (see below). This property makes the Radon transform attractive in rotation invariant vision systems.

## II. GENERAL OBJECTIVE

We consider rotated colour images represented by three matrices in the RGB colour space. The meaning of the images is insignificant in our considerations, and only their visual appearance is taken into account. For the ease of labeling, the images contain animals. We use $64 \times 64$ images, although their size influences only the total learning time. An example

of an image 'baboon' and its version rotated by $30^o$ is given in Fig. 1. The icons derived from all other images used can



Fig. 1.   The 'baboon' image and its version rotated by $30^o$

be seen in Fig. 4.

The network of self-organizing modules is trained to categorize the un-rotated images. The training vectors are based on the Radon transform of the images. The Radon transform is performed for a number of angles, $\theta_i$. After the training is completed, the network will be excited with the rotated copies of the images. We show that if images are rotated by the angles $\theta_i$ used in calculation of the Radon transform, then the categorization is perfect, exactly the same as for the un-rotated images. For all other angles, there is a small categorization error that does not imperil recognition in any way. In addition, the process is very fast, hence might be applicable in human vision. This is really an encouraging result.

Another problem indirectly considered here is related to our quest for a universal "neuronal code". Following our other works, e.g. [22], [23], we demonstrate that for the categorization purposes, the images are internally represented by a set of 3-D values, or labels, conceptually representing the positions of the winners in the first level maps and related post-synaptic excitations. These labels are then used by the second level module producing the final mapping. Although we do not claim biological fidelity of such representation, we nonetheless consider it as a viable conceptual possibility.

### III. RADON TRANSFORM AND ROTATION INVARIANCE

The Radon transform [24] of a continuous function of two real (or one complex) variables $f(z)$, where $z = x + jy$, is the integral transform $Rf(\theta, s)$ consisting of an integral of $f(z)$ over straight lines $z(t) = e^{j\theta}(s + jt)$ (see Fig. 2a).

$$Rf(\theta, s) = \int_{-\infty}^{+\infty} f(e^{j\theta}(s + jt))dt \qquad (1)$$

where $\theta$ is the slope of the line (or ray), $s$ is the distance of the line from the origin, and $t \in (-\infty, +\infty)$ is the line parameter, such that for $t = 0$ the line goes through the point $z(0) = se^{j\theta}$ as indicated in Fig. 2a.

Now, assume that the function $f(z)$ is rotated around the origin by the angle $\alpha$ and calculate the Radon transform of such a rotated function. In the complex domain the rotated
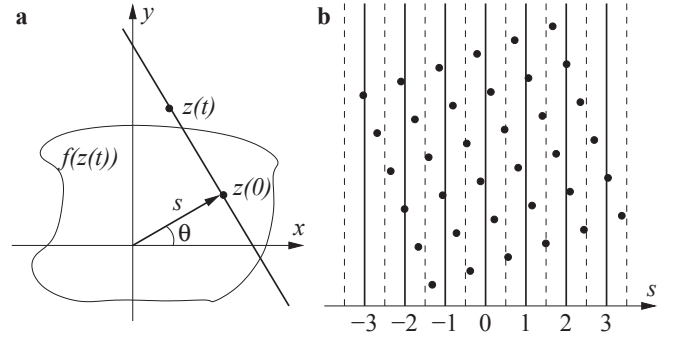


Fig. 2.   The Radon Transform in the continuous (a) and discrete (b) space.

function is $f(ze^{j\alpha})$ and the Radon transform of the rotated function is:

$$Rf(\theta, s; \alpha) = \int_{-\infty}^{+\infty} f(e^{j(\theta+\alpha)}(s + jt))dt = Rf(\theta + \alpha, s) \qquad (2)$$

This simple, but a very powerful relationship says that the Radon transform at the point $(\theta, s)$ of a function rotated by the angle $\alpha$ is equal to the Radon transform of the original, un-rotated function calculated at the point $(\theta + \alpha, s)$. From (2) we can derive some practical methods of calculating the Radon transform specifically suited for discrete space functions like images. With reference to Fig. 2b, where we consider a rotated image-like function in the discrete space, we can note that it might be more accurate to rotate the image (locations of its pixels) rather than the summation lines in the Radon transform. In other words, according to (2), in order to get $Rf(\theta, s)$ we rotate the image by the angle $-\theta$ and calculate the Radon transform for the vertical lines, since, $Rf(\theta - \theta, s) = Rf(0^o, s)$. Hence, in the discrete case, we just sum the values of the pixels located in the strip $(s-0.5, s+0.5]$ without any need for image interpolation.

Now we can consider the central for our considerations problem of forming from the Radon transform $Rf(\theta, s)$ a feature (or signature) function, say, $h(s)$ that can be used to categorize the function $f(x, y)$ in general, and images in particular, in a rotation-invariant fashion. An obvious solution to create such a signature is to integrate the Radon transform $Rf(\theta, s)$ over all angles. This gives the following angle independent result:

$$h(s) = \int_0^{2\pi} Rf(\theta, s)d\theta \qquad (3)$$

Such a feature (signature) function $h(s)$ retains some characteristics of the original function $f(x, y)$, but the angular dependency is removed. That is exactly what we need to achieve the rotational invariance.

In the discrete case, with reference to Fig. 2b, we can observe that the size of the feature vector $\mathbf{h}$ will be equal to the diagonal of the image grid. In the examples that follow, we use $64 \times 64$ images, hence the size of the feature vector $\mathbf{h}$ is equal to $n = \lceil 64\sqrt{2} \rceil = 91$.
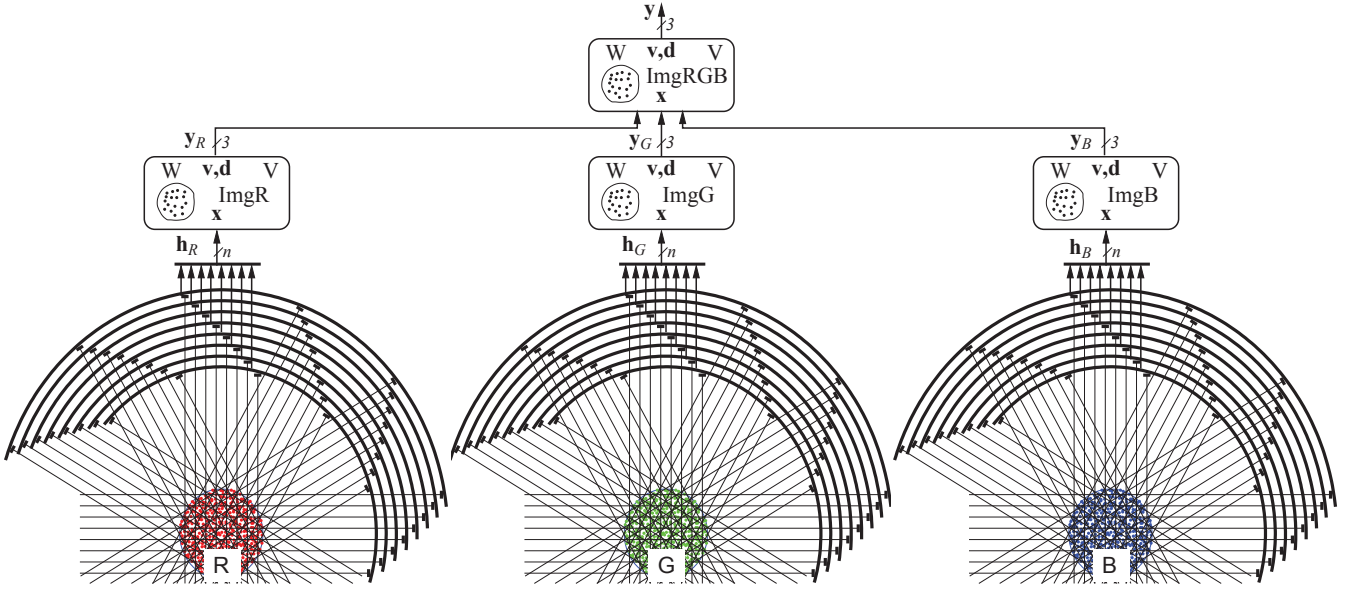
Fig. 3. The structure of the network for rotation invariant categorization of colour visual objects

To conclude the general considerations of this section we observe that general definition of the Radon Transform specifies integration along the set any continuous line. In particular, we could consider a set of circles. This gives an interesting relationship to polar coordinates that needs to be fully investigated. The reader could start with [25] where rotation and scale invariant texture classification based on log-polar wavelet signatures was considered.

## IV. THE STRUCTURE OF THE NETWORK FOR ROTATION INVARIANT CATEGORIZATION OF COLOUR IMAGES

The structure of the network for the rotation invariant categorization of colour image is presented in Fig. 3. At the entry to the network there are three receptive fields marked R, G an B in Fig. 3. The three planes of the rotated, in general, colour images are presented at their respective receptive fields. The second part of the network are summing dendrites webbed over the receptive fields that calculate the Radon transform for each image colour plane. The dendrites are equivalent to the Radon transform rays summing the values of the pixels located at the rays and at their proximity as in (2) and Fig. 2. In the example presented in Fig. 3, for each receptive field we calculate the Radon transform for $m = 6$ angles $\theta$ and $n = 8$ lines located at $s$. In the computational example with images that follows, we use $m = 60$ angles:

$$\theta \in [0^o, 6^o, 12^o, \ldots, 354^o] \tag{4}$$

and $n = 91$ lines located at the integer values of $s$

$$-45 \leq s \leq +45 \tag{5}$$

The next part of the network, above the receptive fields, are circular summing dendrites calculating the feature vectors $\mathbf{h}_R, \mathbf{h}_G, \mathbf{h}_B$ as in (3). It is important to note that these two parts of the network are fundamentally simple, performing

only signal aggregation (summation) which is the most basic biological operation. Therefore, for each image presented at the receptive fields, the three feature vectors are calculated "instantaneously".

The categorization part of the network consists of two-level self-organizing modules. Each colour self-organizing module, e.g. 'ImgR', contains a randomly generated number of neurons, approximately 1000, proportional to the number of images being classified which is 24. The number of neurons is not critical and has not been optimized in any way. Each module, e.g. red, produces a 3-dimensional output $\mathbf{y}_R$ calculated as:

$$\mathbf{y}_R = g(\mathbf{W}_R \cdot \mathbf{h}_R) \tag{6}$$

where $\mathbf{y}_R$ represents output signals from the module, and $g(\cdot)$ describes the Winner-Takes-All function of $\mathbf{W}_R \cdot \mathbf{h}_R$, which produces a 2-D positional vector $\mathbf{v}$ and related postsynaptic activity $d$ forming the output $\mathbf{y}_R = [\mathbf{v}\ d]$. Such 3-D outputs can be thought of as low dimensional signatures, or labels, and can be used in the next hierarchical levels of modules. We postulate that conceptually $\mathbf{y}_R$ is a good candidate for the internal "universal neuronal code". In our specific case, those 3-D labels form a 9-dimensional input vector to the final module 'ImgRGB' that performs the final categorization of images according to:

$$\mathbf{y}_{RGB} = g(\mathbf{W}_{RGB} \cdot \mathbf{x}_{RGB}) \tag{7}$$

where $\mathbf{x}_{RGB} = [\mathbf{y}_R\ \mathbf{y}_G\ \mathbf{y}_B]$ is a 9-D concatenation of the first level outputs.

## V. WORKING OF THE NETWORK

### A. Training

For training we use a set of 24 colour images of the size $64 \times 64$ pixels in the RGB space. An example of such an

image is given in Fig. 1. Each image is then represented by its rotation-invariant feature vectors $\mathbf{h}_R, \mathbf{h}_G, \mathbf{h}_B$, which are used as afferent signals to the respective self-organizing modules. Since the feature vectors are already rotation-invariant, for training we can use either rotated or un-rotated images. Re-using existing software, we normalise the data (images) and the neuronal grid, and project them on the respective unity spheres. This makes it possible to use a simplified 'dot-product' Kohonen learning law [26]. In this case the update of a weight vector $\mathbf{w}_i$ for the $i$th neuron is described by the following expression:

$$\Delta \mathbf{w}_i = \eta \cdot \Lambda_i \cdot (\mathbf{x}^T - d_i \cdot \mathbf{w}_i) \; ; \quad d_i = \mathbf{w}_i \cdot \mathbf{x} \qquad (8)$$

where $\Lambda_i$ is the neighbourhood function, Gaussian in our case, centered at the position of the winning neuron, and $d_i$ is the **post-synaptic activity** of the $i$th neuron. The computational simplification is that the distances between vectors are calculated by their dot-product, which is equal to the cosine of the angle between them. The result of learning are four



Categorization of the rotated images

30,60,90,120,150,180,210,240,270,300,330,[deg]

Fig. 5. The result of testing categorization of visual objects rotated by a subset of angles used in the Radon transform



Fig. 4. Three first-level R, G, B maps and the final map representing categorization of colour visual object

maps describing locations of images on a neuronal grid as in Fig. 4. There are three first level maps, for the R, G and B components of the colour images, and the final map incorporating information from the constituent maps as in (7).

### B. Testing 1

After training we test the working of the network when rotated images are applied to the receptive fields. Firstly, we rotate images by the subset of 10 angles from (4), namely, $\alpha \in 30^o, 60^o, \dots, 330^o$. The results of categorization of these rotated images are perfect, as presented in Fig. 5. Instead of icons as in Fig. 4, for clarity we use the names of the visual
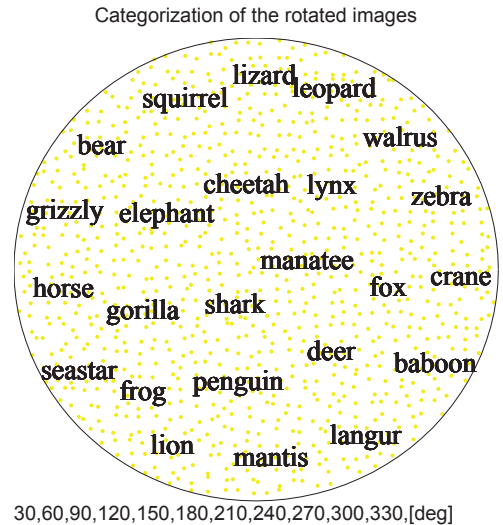
objects located at the position of the winner. What we see in Fig. 5 are names superimposed on each other 11 times. This is really an amazingly "boring" result, as predicted by (3) and subsequent considerations. The rotated objects are categorized exactly as un-rotated ones.

We note, that we can set up the number of the Radon Transform angles used in training to the required accuracy.

### C. Testing 2

It is interesting to test the behaviour of the network when it 'sees' images rotated by angles different than those used to calculate the Radon transform and the resulting feature vectors $\mathbf{h}_R, \mathbf{h}_G, \mathbf{h}_B$. In this test we use rotation angles $\alpha = -18^o, -16^o, \dots, 16^o, 18^o$. The results of applying such images to the network are presented in Fig. 6a,b,c. Inspecting the mapping in Fig. 6a we can see that for some images there is a shift in position of the winner. This mapping error is shown in Fig. 6b. The numbers show the average shift of the position of the winner for the rotated images. It can be seen, that "perceptually", the rotated images are well inside the clusters related to a specific stimulus, see Fig. 4. To put some measure against the perception, in Fig. 6c we show the confidence level measured by the postsynaptic activity in module 'ImgRGB' (see Fig. 3). For the perfect "perception", the postsynaptic activity, measured by the inner product of the stimulus and the respective weight vector, is equal to unity. It drops a bit for the angles that were not included in the Radon transform as shown in Fig. 6c. The maximum equivalent angular error between the stimulus and the respective weight vector is $\arccos(0.9999) < 1^o$. The result demonstrates that the network can easily generalized for the angles not included in the Radon transform.

### VI. Conclusion

We have presented a network of self-organized modules that can correctly categorize colour objects in a rotation invariant

**a**

Categorization of the rotated images

−18,−16,−14,−12,−10,−8,−6,−4,−2,0,2,4,6,8,10,12,14,16,18,[c

**b**

Average location error for rotated images
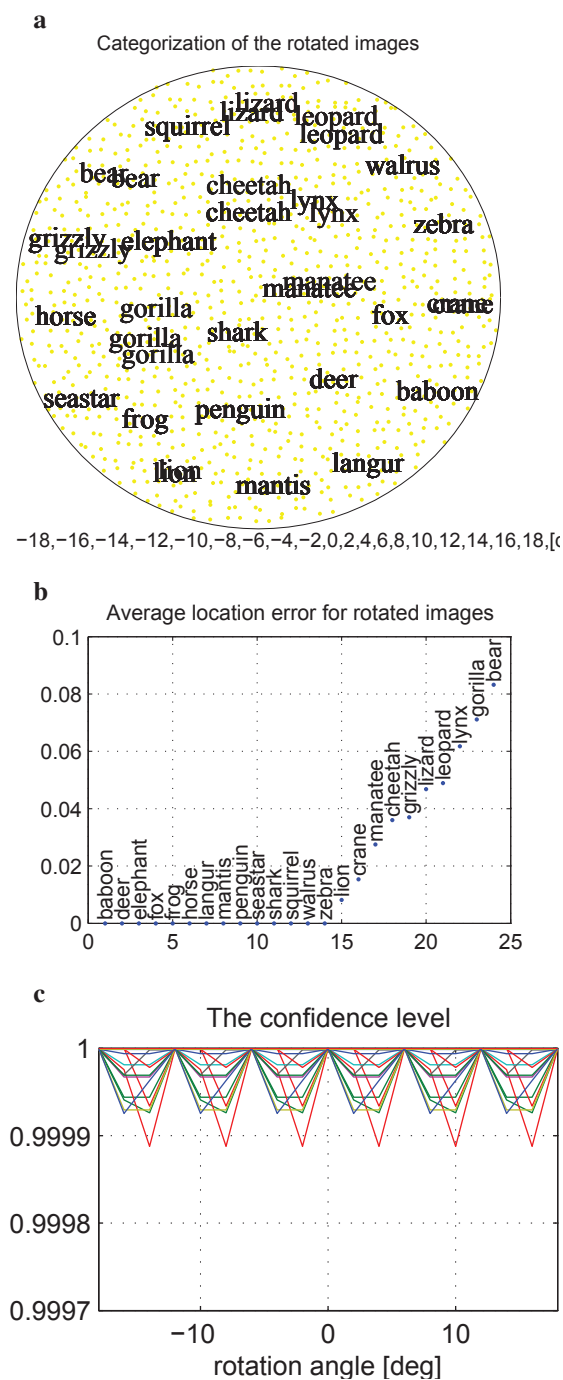
**c**

The confidence level

Fig. 6. The result of testing categorization of images rotated by angles not included in the Radon transform: (a) The mapping of the rotated images, (b) an average mapping error, (c) The confidence level

fashion. At the foundation of the solution is a novel rotation invariant feature vector, which in the continuous space is formed as the angular integral of the Radon transform of a function of two variables as in (3). In the case of colour images three Radon transforms and related feature vectors are calculated. These feature vectors are used to categorize colour images in a network of 3+1 self-organizing modules. The 3-D outputs from the first layer modules are considered to be labels

used by the higher-level self-organizing modules and can be thought of as a "universal neuronal code". In our specific case three 3-D labels from the first layer form the 9-D afferent vector for the second layer. The network is trained for un-rotated images and then tested for rotated images. It has been demonstrated that rotation of images by the angles included in calculation of the Radon transform results in the perfect categorization. For the angles in between, that is, those not included in the Radon transform a small shift in categorization might occur, keeping the objects, however, well inside their respective clusters. Since calculation of the rotation-invariant feature vectors is very simple and involves only summations of signals (pixel values), hence very fast, it is postulated that such a mechanism might be included in the biological vision systems.

REFERENCES

[1] A. P. Papliński, "Rotation invariant categorization of visual objects using Radon transform and self-organizing modules," in *Lect. Notes in Comp. Sci.*, vol. 6444. Springer, 2010, pp. 360–366.

[2] O. Al-Shaykh and J. Doherty, "Invariant image analysis based on Radon transform and SVD," *IEEE Trans. Circ. Syst. II*, vol. 42, no. 2, pp. 123–133, 1996.

[3] N. Kingsbury, "Rotation-invariant local feature matching with complex wavelets," in *Proc. 14th Europ. Sig. Proc. Conf.*, Florence, Italy, 2006, pp. 1–5.

[4] J. Fehr, "Local rotation invariant patch descriptors for 3D vector fields," in *Proc. Int Conf. Pat. Rec*, Istanbul, 2010, pp. 1381–1384.

[5] J. D. B. Nelson and N. G. Kingsbury, "Enhanced shift and scale tolerance for rotation invariant polar matching with dual-tree wavelets," *IEEE Trans. Img. Proc*, vol. 20, no. 3, pp. 814–821, 2011.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comp. Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[7] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. Conf. Comp. Vis. Pat. Rec.*, Washington, 2004, pp. 511–517.

[8] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. 9th Europ. Conf. Comp. Vis.*, vol. LNCS 3951. Graz, Austria: Springer, 2006, pp. 404–417.

[9] L. Juan and O. Gwun, "A comparison of SIFT, PCA-SIFT and SURF," *Int. J. Img. Proc.*, vol. 3, no. 4, pp. 143–152, 2009.

[10] O. C. Hamsici and A. M. Martinez, "Rotation invariant kernels and their application to shape analysis," *IEEE Tran. PAMI*, vol. 31, no. 11, pp. 1985–1999, 2009.

[11] E. Magli, L. L. Presti, and G. Olmo, "A pattern detection and compression algorithm based on the joint wavelet and Radon transform," in *Proc. IEEE 13th Int. Conf. Dig. Sig. Proc.*, 1997, pp. 559–562.

[12] A. Warrick and P. A. Delaney, "Detection of linear features using a localized Radon transform with a wavelet filter," in *Proc. ICASSP*, 1997, pp. 2769–2772.

[13] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms," *IEEE Trans. Img. Proc*, vol. 14, no. 6, pp. 783–795, 2005.

[14] G. Yu, W. Cao, and Z. Li, "Rotation and scale invariant for texture analysis based on Radon transform and wavelet transform," in *Proc. 3rd ICPCA*, 2008, pp. 704–708.

[15] W. Yao and C.-M. Pun, "Invariant shape representation by Radon and wavelet transforms for complex inner shapes," in *Proc. IEEE Int. Conf. Inform. Autom.*, 2009, pp. 1144–1149.

[16] S.-S. Xiao and Y.-X. Wu, "Rotation-invariant texture analysis using Radon and Fourier transforms," *J. Phys.: Conf. Ser.*, vol. 48, pp. 1459–1464, 2007.

[17] G. Chen and B. Kégl, "Feature extraction using Radon, wavelet and Fourier transform," in *Proc. IEEE Int. Conf. Syst. Man and Cybernetics*, 2007, pp. 1020–1025.

[18] G. Liu, Z. Lin, and Y. Yu, "Radon representation-based feature descriptor for texture classification," *IEEE Trans. Img. Proc*, vol. 18, no. 5, pp. 921–928, 2009.

[19] M. Miciak, "Character recognition using Radon transformation and principal component analysis in postal applications," in *Proc. Int. Multiconf. Comp. Sci. Info. Tech.*, 2008, pp. 495–500.

[20] M. Hejazi, G. Shevlyakov, and Y.-S. Ho, "Modified discrete Radon transforms and their application to rotation-invariant image analysis," in *Proc. IEEE Workshop Mult. Sig. Proc*, 2006, pp. 429–434.

[21] F. Hjouj and D. W. Kammler, "Identification of reflected, scaled, translated, and rotated objects from their Radon projections," *IEEE Trans. Img. Proc*, vol. 17, no. 3, pp. 301–310, 2008.

[22] A. P. Papliński, L. Gustafsson, and W. M. Mount, "A recurrent multi-modal network for binding written words and sensory-based semantics into concepts," in *ICONIP 2011, LNCSI*, B.-L. Lu, L. Zhang, and J. Kwok, Eds., vol. 7062.   Springer, 2011, pp. 413–422.

[23] T. Jantvik, L. Gustafsson, and A. P. Papliński, "A self-organized artificial neural network architecture for sensory integration with applications to letter–phoneme integration," *Neural Computation*, pp. 1–39, 2011, doi:10.1162/NECO_a_00149.

[24] "Radon transform," http://en.wikipedia.org/wiki/Radon_transform.

[25] C.-M. Pun and M.-C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," *IEEE Trans. PAMI*, vol. 25, no. 5, pp. 590–603, 2003.

[26] T. Kohonen, *Self-Organising Maps*, 3rd ed.   Berlin: Springer-Verlag, 2001.