Image Object Search Combining Color with Gabor Wavelet Shape Descriptors

Darryl D. Anderson and Mark S. Drew

School of Computing Science, Simon Fraser University,
Vancouver, B.C. Canada V5A 1S6

ABSTRACT

An image and object search and retrieval algorithm is devised that combines color and spatial information. Spatial characteristics are described in terms of Wiskott's jets formulation, based on a set of Gabor wavelet functions at varying scales, orientations and locations. Color information is first converted to a form more impervious to illumination color change, reduced to 2D, and encoded in a histogram based on a new stretched chromaticity space for which all bins are populated. An image database of 27,380 images is devised by replicating 2,738 JPEG images by a set of transforms that include resizing, various cropping attacks, JPEG quality changes, aspect ratio alteration, and reducing color to grayscale. Correlation of the complete encode vector is used as the similarity measure. For both searches with the original image as probe within the complete dataset, and with the altered images as probes with the original dataset, the grayscale, the stretched, and the resized images had near-perfect results. The most formidable challenge was found to be images that were cropped both horizontally as well as vertically. The algorithm's ability to identify *objects*, as opposed to just images, is tested. In searching for images in a set of 4 classifications, the jets were found to contribute most analytic power when objects with distinctive spatial characteristics were the target.

Keywords: Image Classification, Content-based Image Retrieval, Object Recognition, Gabor Wavelet Filters, Chromaticity Histogram, Illuminant Invariance

1. INTRODUCTION

Image content-based search and retrieval has the potential to be at least as useful, if not more so, than traditional text-based searching. Increases in processing power, bandwidth and storage capability have increased the availability of multimedia data. These collections of multimedia data need to be organized based on content. As image processing is typically computationally expensive, the need for efficient and scalable algorithms to retrieve image content is apparent. An efficient approach is to analyze an image and generate a signature based on distinguishing information. Images are then correlated based on their signatures. This method provides a fast and scalable method of image recognition, because the signature is generated offline as a preprocessing step and stored in a database. Typical image and object recognition algorithms analyze color and shape information.

A traditional color-based object recognition approach uses color histogram information to compare objects. Though color information is a powerful indicator in object recognition, it has several difficulties. First, this approach discards information about the objects' spatial properties, which is another powerful indicator of object similarity. It is very likely that two different objects could have similar color decompositions and distinctly different shapes, resulting in an incorrect false positive identification. Second, color information varies under different lighting conditions and with different cameras. Although the colors of the objects remain the same, the colors captured by the camera can vary dramatically. Fortunately, there has been significant progress in the field of color invariance to overcome this difficulty. In this study, an illumination-invariant chromaticity method based on work by Drew and Au[1] is used to provide the color analysis.

Shape-based object recognition compares objects based on their measurable shape. One approach is to generate a zero crossing signature. The second derivative of the object's edge is used to reveal the number

Further author information: (Send correspondence to D.D.A.)
D.D.A.: E-mail: dda@cs.sfu.ca
M.S.D.: E-mail: mark@cs.sfu.ca

and placement of the zero crossings (where the second derivative changes sign). Another approach is to infer the texture or 3D shape from an image. This is typically done with a form of frequency analysis. Shape-based object recognition has at least one major drawback. Objects that have a complex 3D shape yield a very different shape analysis with only slight changes in rotation, pitch or yaw. In this study, a variation of Wiskott's Gabor wavelet filter based Jets[2] are used to provide the frequency analysis.

The focus of this paper is the improvement of current image recognition by using color analysis combined with frequency analysis. Color and shape indicators are used to recognize images, thus improving on the results of either method used individually. Experiments were conducted that show the effectiveness of the method produced at content-based image retrieval, indexing and recognizing images and object recognition.

2. BACKGROUND

2.1. Linear Correlation

Linear correlation is the most widely used measure of association between variables that are ordinal or continuos, rather than nominal.[3] Given two arrays, x and y of length N, having pairs of quantities (x_i, y_i), $i = 1, ..., N$, the linear correlation coefficient r is given by the formula:

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})^2}} \tag{1}$$

where $\bar{x}$ is the mean of the x_i and $\bar{y}$ is the mean of the y_i.

The value of r is in the range [-1,1]. Complete positive correlation, with $r = 1$, occurs when the data values lie on a perfectly straight line with a positive slope (with x and y increasing together). Similarly, a value of -1, complete negative correlation, occurs when the data values lie on a perfectly straight line with a negative slope (with x and y decreasing together). The closer r is to zero, the more uncorrelated the data values are.

2.2. Chromaticity Histograms

Research on image content indexing and retrieval started by focusing on extraction and correlation of global image feature vectors. In one of the early works on image retrieval, Swain and Ballard[4] used *histogram intersection* to correlate color histograms of two images. First, a color histogram $\mathbf{H}_i$ is generated for each image i in the database. The histogram is then normalized, and stored in the database. For a model image from the database, its histogram $\mathbf{H}_m$ is intersected with all database image histograms $\mathbf{H}_i$ according to the equation:

$$\sum_{j=1}^{n} \min(H_i^j, H_m^j), \tag{2}$$

where superscript j denotes histogram bin j, and each histogram has n bins. The closer the intersection value is to 1, the better the images match. Computing the intersection value is fast, but it is sensitive to color quantization.

Chromaticity histograms have been shown to be an improvement on color histograms.[5] There are two advantages of using a chromaticity color space. First, chromaticity color space reduces the dimensionality of color to 2, which among other things significantly reduces the size of the histogram. Second, because chromaticity is a ratio of color bands, it has the effect of removing shading,[1] which contributes to illumination invariance. Linear chromaticity is calculated using Equation 3.

$$(r, g) = (R, G)/(R + G + B) \tag{3}$$

The calculation of the linear chromaticity introduces a problem. As ratios are not evenly distributed, they do not fully utilize the evenly spaced histogram bins. Drew and Au[1] point out that because linear chromaticity obeys $r+g \leq 1$, there exists a straight diagonal edge in a chromaticity space histogram. To overcome the negative effects of the diagonal edge, Drew and Au used a spherical chromaticity space of the form described in Equation

4, to mitigate a ringing effect in the Fourier domain caused by the diagonal edge. Spherical chromaticity does not eliminate the diagonal edge, but improves matters by replacing it with a circular edge. In addition, spherical chromaticity space improved upon the linear model by utilizing more of the histogram bins.

$$(r, g) = (R, G)/\sqrt[p]{R^p + G^p + B^p} \tag{4}$$

Here we propose the use of a 2D *stretched* chromaticity space similar to that in[6] * shown in Equation 5, which utilizes all of the histogram bins and eliminates the edge effect altogether.

$$\begin{aligned} (r, g) &= (R, G)/(R + G + B) \\ r &= \begin{cases} r + g & : \quad if\ r \geq g \\ 2r & : \quad otherwise \end{cases} \\ g &= \begin{cases} 2g & : \quad if\ r \geq g \\ r + g & : \quad otherwise \end{cases} \end{aligned} \tag{5}$$

2.3. Gabor Wavelet Filter

The use of the 2D Gabor filter in computer vision was introduced by Daugman in the late eighties.[7,8] Since that time it has been used in many computer vision applications including image compression,[8] edge detection,[9] texture analysis,[10] object recognition[11] and facial recognition.[2,12–15]

The general form for a complex-valued 2D Gabor function is a planar wave attenuated by a Gaussian envelope:

$$\Psi(k, x) = \frac{k^2}{\sigma^2} exp(-\frac{k^2 x^2}{2\sigma^2})[exp(ik.x) - exp(-\frac{\sigma^2}{2})] \tag{6}$$

In order to render the filters insensitive to the overall level of illumination, the term $exp(-\frac{\sigma^2}{2})$ is subtracted. The multiplicative factor k^2 ensures that filters tuned to different spatial frequency bands have approximately equal energies.

2.4. Jets

Wiskott uses jets extensively in his facial recognition and scene analysis applications.[2,14–16] Jets are a grouping of wavelets in varying orientations and frequencies evaluated at a single point. A jet is a condensed and robust representation of a local grey value distribution, termed a *local expert.*[2] Wiskott further describes a jet as being based on a Gabor wavelet transform, which is a convolution with a family of complex Gabor wavelets having the shape of plane waves restricted by a Gaussian envelope function. The wavelets are similar in the sense that they can all be generated from a mother wavelet by rotation and scaling. All complex coefficients of the transform taken at one image location form a jet.

A small displacement may lead to very different coefficients. However, the magnitudes vary slowly and can be used directly for comparison.[2] Wiskott also mentions that jets are robust with respect to illumination variations, scaling, translation and distortion. Another benefit to Gabor wavelets is that they are a good model for the receptive fields of complex cells in primary visual cortex of primates.[2]

---

*This chromaticity space is like that in,[6] but properly produces the range {[0..1],[0..1]}, rather than {[0..1],[0..2]}.

2.5. Gabor Filter Frequencies

When performing a frequency analysis it is important to consider what frequencies should be evaluated. Low frequency information remains more stable across images. Nestares et al.[17] selected $\frac{Nyquist}{2}$ as the highest central frequency of their implemented Gabor filter banks. In accordance with Nyquist's theorem, lower resolution images have fewer useful frequencies that can be used. High frequency information has been shown to be a differentiating factor in texture analysis in high-resolution images,[10] but the majority of images are not high resolution. In addition, high frequency information in images is often associated with edges and noise[18] and we are not directly concerned with edge information. For this implementation, we therefore use only the low frequency information.

3. PROCEDURE

3.1. Frequency Analysis

For this application, a variation on Wiskott's jet was used that evaluated the Gabor filter in three different directions. The three directions were at 0 degrees (horizontal), at 45 degrees and at 90 degrees (vertical). There are two sizes of jets that are evaluated. The small jets evaluate the lowest frequency and the bigger jets evaluate the two lowest frequencies.

The jet uses σ_x and σ_y to determine the Gaussian width in the x and y directions. In many applications, these two values are equal, and hence yield a circular filter. They were used to produce elliptical shaped filters, *with the same aspect ratio* as the images on which they were evaluated. Elliptical filters are used to analyze foreshortened objects or texture patterns. Images that have been stretched, or have otherwise had the aspect ratio altered, are analogous to a foreshortened view. Binding the shape of the filter to the aspect ratio of the image produces an algorithm that is invariant to aspect changes.

The $exp(-\frac{\sigma^2}{2})$ factor, listed in Equation 6, was not included. This is usually subtracted to yield a zero DC response, since the integral of the cosine is always larger than the sine. This makes the response magnitudes invariant under different lighting contrast conditions. Since the outer ends are severely attenuated by the Gaussian, it will have little effect on the overall response.

The jets must be evaluated at locations that will provide distinguishing information. As image content is not known *a priori*, locations covering the image extents were sampled to gather enough distinguishing information. Four considerations were taken into account when devising this method:

1. As the content of the image was not known beforehand, no assumptions could be made as to the placement or orientation of the distinguishing content.
2. The wavelet placement must be predictable and repeatable.
3. The wavelet placement method must be size invariant as the image dimensions can vary.
4. The wavelet placement must be invariant to aspect changes in the imagery.

To achieve goal 1, two uniform grids of Gabor filters were used over the image. The first layer of jets was a uniform 10x10 grid. The second layer of jets was a uniform 9x9 grid that was twice the size of the jets in the first layer and overlapped the same areas. Because of their increased size, two frequencies were evaluated at each location. To achieve goal 2, the size and placement of the filters were made a function of the image extents. In this way, the filter size and shape were proportional to the size and shape of the image, satisfying goals 3 and 4.

3.2. Color Analysis and Illumination Invariance

The color analysis was performed using a method similar to a method used by Drew and Au.[1] As with their method, the color channels were normalized before moving into chromaticity space. Normalization of the color channels greatly attenuates dependence on both luminance and lighting color.[1] This is accomplished by first dividing each color channel by its mean, then normalizing each pixel to length 1 by dividing by the square root of the sum of the squares of the RGB values. Iterating in this manner has been shown to provide convergence after five iterations[19,20]; however good results can still be achieved in one iteration with far less computation.

A 2D stretched-chromaticity histogram was created by breaking each band into 32 equal sections. Histogram bin counts were determined and then normalized by making them a percentage of the total number of pixels. The histogram values were then appended to the filter coefficients to complete the encode vector. Linear correlation was used for correlating the histogram values. Like Swain and Ballard's method, this method is fast, but has the advantage of being less sensitive to color quantization. When dealing with grayscale images, color information was ignored.

3.3. Image Encoding

The image was encoded in the following manner:

1. A Gabor wavelet filter was used in multiple locations over the image. The coefficients, produced by evaluating the Gabor filters at each location, were added to an encode vector.
2. A color decomposition was performed by adding every pixel to the appropriate 2D stretched-chromaticity histogram bin. The bins counts were normalized and added to the encode vector.

The data from the frequency analysis and chromaticity histogram was combined to form an encode vector which acted as a signature. The similarity between any two encode vectors was calculated using linear correlation. Linear correlation produces a number in the range of [-1,1]. The absolute value of the correlation was used to rank the images — this is valid because it is the distance from zero that represents correlation.

The descriptors produced by the Gabor filter are always positive. Moreover, the values in the chromaticity histogram were normalized by being calculated as a percentage of the total number of pixels and then cast into the range $[0, 2^{16}]$. Because the values in the encoded vector are always positive, the results were stored in an unsigned 16-bit integer. The small data type was used to increase the speed of the correlation and to reduce the data that was stored in the database. If the results of the Gabor filter were greater than 2^{16} then the result was truncated and the coefficient was assigned to be 2^{16}. In the dataset used for this experiment this condition never occurred.

4. EXPERIMENT

An image data set was assembled for testing the image recognition algorithm. 2,738 JPEG images of varying content were chosen from a photo library † of over 40,000 images. 2,698 of those JPEG images were selected by choosing five images, if available, from each directory within the photo library. The purpose of these images was to obtain varying content in which the image recognition capabilities could be tested. The remaining 40 JPEG images were selected by hand for their similarity to each other and images already present in the dataset. The purpose of these 46 images (6 were already present in the dataset) was to obtain content in which the similar image finding capabilities could be tested as well as testing the object recognition capabilities.

In order to test the robustness of various aspects of the algorithm, 10 variations of each image were used. The format and motivation behind the variations are listed in Table 1. All the original images and their variations were loaded into a database along with the encode vectors that were produced from those images. With ten variations of each image, the total number of images in the dataset was 27,380.

Testing began by isolating each image variation in the dataset, so that each search was restricted to the 2,738 images of the same variation. Using the original image as the probe, each image variation was searched

†Corel Gallery Photo Library.

Table 1. Format and motivation behind the image copies included in the test dataset.

Format of Image Copy	Purpose
Original JPG image	Original
Grayscale image	Color invariance
70% resized image	Invariance to small size change
30% resized image	Invariance to large size change
24% height cropped image	Spatial shift invariance
24% width cropped image	Spatial shift invariance
24% height and width cropped image	Spatial shift invariance
20% quality JPG image	Pixelation and color invariance
30% width stretched image	Aspect change invariance
30% height stretched image	Aspect change invariance

for in turn. This allowed the system's performance to be measured under each of the conditions produced by the image variations. The top fifty matches from each search were analyzed. The search results were used to calculate the cumulative match characteristic (CMC) score. The CMC score is the cumulative count of the correct number of returns. It is shown as a percentage of the total number of correct images expected. A score of 80% achieved by looking at only the first return from each search indicates that 80% of the images returned the correct image variation in the top ranked position. The CMC scores for each image variation search are summarized in Table 2. The system found the Original, Greyscale and 20% quality images with 100% accuracy in the top position.

Table 2. Using every original JPEG image, each image variation was searched for in turn and its CMC score was recorded. The percentage of the variations returned within each ranking category is listed.

Image Variation Searched For	% Returned in the Top 1	% Returned in the Top 2	% Returned in the Top 5	% Returned in the Top 10	% Returned in the Top 25
Original	100	100	100	100	100
Grayscale	100	100	100	100	100
70% resized	99.71	99.74	99.89	99.89	99.93
30% resized	83.16	84.88	87.98	89.92	92.44
24% Y crop	81.81	84.44	87.66	89.70	92.18
24% X crop	76.84	79.80	84.00	86.78	90.21
24% Y & X crop	46.93	52.92	61.03	67.79	76.15
20% quality	100	100	100	100	100
30% X strech	92.73	93.54	94.67	95.51	96.49
30% Y strech	99.71	99.71	99.96	99.96	99.96

The dataset was then restricted to the 2,738 original images. All of the image variations were then used as probes to search for the original images. The CMC scores generated by using each image variation as the probe image are summarized in Table 3. The results indicate similar relative strengths in the system. Cropping in both directions posed the biggest problem, requiring 25 returns before a CMC score of 82% was achieved. Cropping in the X and Y directions had the correct returns in the first position between 80% and 83% respectively. The greyscale, stretched and resized images had near-perfect results with CMC scores ranging from 92% to 99% in the top search return. Original and 20% quality images still generated perfect results.

Table 3. Using every image variation, each original JPEG was searched for in turn and its CMC score was recorded. The percentage of the original images returned within each ranking catagory is listed.

Image Variation	% Returned in the Top 1	% Returned in the Top 2	% Returned in the Top 5	% Returned in the Top 10	% Returned in the Top 25
Original	100	100	100	100	100
Grayscale	99.96	99.96	99.96	99.96	100
70% resized	99.67	99.74	99.82	99.82	99.82
30% resized	92.44	93.50	94.67	95.40	96.38
24% Y crop	83.89	87.03	90.14	92.26	94.85
24% X crop	80.13	83.27	86.23	88.75	91.75
24% Y & X crop	50.29	57.71	66.44	74.03	82.03
20% quality	100	100	100	100	100
30% X strech	92.55	93.83	95.36	96.31	97.81
30% Y strech	99.67	99.85	99.96	99.96	100

In order to test the overall capabilities of the system, searches were performed using all image variations as probes against the entire dataset. The number of correct returns was counted within the top 10, 15, 25 and 50 returns. With ten copies of each image, the ultimate goal is to have all ten copies within the top ten returns. With 2,738 images of each variation and ten correct returns for each of them, the highest score possible is 27,380. The score, as a percentage of the total, for each image variation search is summarized in Table 4.

Table 4. Using every image variation, every other image variation was searched for and the percentage of correct returns found is listed.

Image Type	% of correct ret. in the top 10	% of correct ret. in the top 15	% of correct ret. in the top 25	% of correct ret. in the top 50
Original	90.24%	92.44%	94.09%	95.56%
Grayscale	60.24%	61.32%	62.45%	64.03%
70% resized	89.18%	91.58%	93.24%	94.93%
30% resized	79.76%	83.05%	85.77%	88.26%
24% Y crop	72.57%	76.63%	80.19%	84.14%
24% X crop	70.20%	73.72%	77.13%	80.81%
24% Y & X crop	60.63%	65.39%	70.28%	75.43%
20% quality	90.24%	92.44%	94.09%	95.56%
30% X stretch	76.34%	80.64%	84.15%	87.83%
30% Y stretch	87.68%	90.30%	92.28%	94.02%

Out of the 2,738 images, 46 were used for *similar image testing*. The images were divided into four classifications. The four visually similar classifications that were chosen were Wedding, Abstract Color, Mug shots and Parisian Door. Fig. 1 shows example images from these categories.

Figure 1. Sample images from 4 categories: Wedding, Abstract Color, Mug-shot, Parisian_Door

The similar images in each of the classifications were counted. There were 6 Wedding images, 12 Abstract color images, 13 Mug-shot images and 15 Parisian Door images. Table 5 summarizes the results of searching for visual similar images using only one of the images as the probe. For this experiment, only the original images were included in the searches. The number of similar scenes in each classification is listed in the second column of Table 5. The third column is the number of similar images returned in the top five. The fourth column is the number of similar images that were returned in the top ten. The fifth column is the number of similar images that were returned in the top fifteen returns. The sixth column is the percentage of similar images that were returned out of the total, up to a maximum of 15. The top five returns were completely accurate using this method. The percentage of similar scenes in the top fifteen ranged from 75% to 100%.

Table 5. Using 46 images in four classifications of similar images, the results of searching for similar image content is listed.

Classification	Number of sim. scenes	Number of top 5 sim.	Number of top 10 sim.	Number of top 15 sim.	% of total up to 15
Wedding	6	5	6	6	100%
Parisian Door	15	5	10	14	93%
Mug-shots	13	5	9	12	92%
Abstract Color	12	5	8	9	75%

The similar image testing also reveals the systems applicability to object recognition. The 46 images are essentially pictures of objects, in particular the Parisian Doors. The dataset contained 15 different doors. All of the doors had some similar qualities. All were inset in brick buildings and occupied the same proportion of the image. However, there were significant differences as well. Some of the doors were rounded while some were square and some had elaborate detailing while others were quite simple. Still the system was able to pull the images of the doors to the top of the returns. In this respect, there is significant potential for the application of this algorithm in object recognition.

5. CONCLUSIONS AND FUTURE WORK

This project was conducted as a proof of concept for a hybridized approach to object recognition. Based on the results of this project, there is a significant potential for this algorithm. Recognition rates for image recognition were near perfect, 92% or higher, for all image variations except cropping and severe resizing which ranged from

47% to 84%. Similar image finding also produced encouraging results, with similar scenes being returned with between 75% and 100% success.

ACKNOWLEDGMENTS

One of the authors (DDA) acknowledges the support of the Science Council of British Columbia through its Graduate Engineering And Technology (GREAT) scholarship. This research has also been supported in part by Imagis Technologies as the industrial sponsor for the GREAT scholarship. The other author (MSD) is also indebted to the Natural Sciences and Engineering Research Council of Canada for a Discovery Research Grant.

REFERENCES

1. M. Drew and J. Au, "Video keyframe production by efficient clustering of compressed chromaticity signatures," in *ACM Multimedia '00*, pp. 365–368, 2000. http://www.cs.sfu.ca/∼mark/ftp/AcmMM00.
2. L. Wiskott, "Phantom faces for face analysis," *Pattern Recognition* **30**(6), pp. 837–846, 1997.
3. W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipies in C: The Art of Scientific Computing*, Cambridge University Press, New York, New York, USA, 1988.
4. M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision* **7**(1), pp. 11–13, 1991.
5. M. Drew, J. Wei, and Z. Li, "Illumination–invariant color object recognition via compressed chromaticity histograms of color–channel–normalized images," in *ICCV98*, pp. 533–540, IEEE, 1998.
6. D. Paulus, K. Horecki, and K. Wojciechowski, "Localization of colored objects," in *International Conference on Image Processing: ICIP '00*, pp. Vol III 492–495, 2000.
7. J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by twodimensional visual cortical filters," *Journal of the Optical Society of America* **2**(A), pp. 1160–1169, 1985.
8. J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. on Acoustics Speech and Signal Processing* **36**(7), pp. 1169–1179, 1988.
9. R. Mehrotra, K. Namuduri, and N. Ranganathatn, "Gabor filter based edge detection," *Pattern Recognition* **25**(12), pp. 1479–1494, 1992.
10. A. Monadjemi, B. T. Thomas, and M. Mirmehdi, "Experiments on high resolution images towards outdoor scene classification," tech. rep., University of Bristol, Department of Computer Science, February 2002.
11. M. Lades, J. Vorbrggen, J. Buhmann, J. Lange, C. von der Malsburg, R. Wrtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on Computers* **42**, pp. 300–311, March 1993.
12. M. Lyons, J. Budynek, A. Plante, and S. Akamatsu, "Classifying facial attributes using a 2-d gabor wavelet representation and discriminant analysis," in *Proceedings of the 4th International Conference on Automatic Face and Gesture Recognition*, pp. 202–207, March 2000.
13. M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," in *Proceedings, Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200–205, April 1998.
14. L. Wiskott, "The role of topographical constraints in face recognition," *Pattern Recognition Letters* **20**(1), pp. 89–96, 1999.
15. L. Wiskott, *Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis*, vol. 53 of *Reihe Physik*, Verlag Harri Deutsch, Thun, Frankfurt am Main, Germany, 1995. (PhD thesis).
16. L. Wiskott, "Segmentation from motion: Combining Gabor- and Mallat-wavelets to overcome the aperture and correspondence problems.," *Pattern Recognition* **32**(10), pp. 1751–1766, 1999.
17. O. Nestares, R. Navarro, J. Portilla, and A. Tabernero, "Efficient spatial-domain implementation of a multiscale image representation based on gabor functions," *Journal of Electronic Imaging* **7**(1), pp. 166–173, 1998.
18. M. Sonka, V. Hlavac, and R. Boyle, *Image Processing Analysis and Machine Vision*, International Thomson Computer Press, London, 1996.

19. G. Finlayson and S. Hordley, "Color constancy at a pixel," *J. Opt. Soc. Am. A* **18**, pp. 253–264, Feb. 2001. Also, UK Patent application no. 0000682.5. Under review, British Patent Office.
20. G. Finlayson, S. Hordley, and M. Drew, "Removing shadows from images," in *ECCV 2002: European Conference on Computer Vision*, pp. 4:823–836, 2002. Lecture Notes in Computer Science Vol. 2353, http://www.cs.sfu.ca/∼mark/ftp/Eccv02/shadowless.pdf.