Image Retrieval Using Colour Co-occurrence Histograms

Linjiang Yu and Georgy Gimel'farb
CITR, Department of Computer Science, Tamaki Campus,
The University of Auckland , New Zealand
lyu011@ec.auckland.ac.nz, g.gimelfarb@auckland.ac.nz

Abstract

Content-based image retrieval (CBIR) has been intensively studied recent years due to its importance in various database management and computer vision applications. Searching by an image example that allows to retrieve a given image or similar images from a large image collection is one of the most challenging CBIR problems today. The paper proposes and investigates a new algorithm for a partial solution of this problem. The algorithm uses combined colour – texture features to find out whether an image contains spatially homogeneous colour textured regions similar to the given example (training image). First, quantization of the HSV colour space focuses only on the colours to be found during the search. Secondly, similarity between characteristic normalised colour co-occurrence histograms (nCCHs) in the moving windows over the image and the like training nCCHs is measured to detect the desired regions. Finally, the frequency distributions of the similarity values are compared to rank the images in the database in their similarity to the training image. Our experiments show that the proposed algorithm effectively retrieves images containing the desired textures.

Keywords: content-based image retrieval (CBIR), texture, colour co-occurrence histogram (CCH), pairwise interaction

1 Introduction

Image retrieval has important practical applications in database management and computer vision (see, for instance, comprehensive surveys in [1, 2, 3]). An effective image retrieval system should integrate both text-based [4, 5] and content-based image retrieval (CBIR) techniques. Limitations of the former ones are the large amount of human's labour for manual image annotation and subjectivity of human perception of rich image contents. The latter techniques try to overcome these limitations by relating image content to particular quantitative image features such as colour, texture, shape of objects, and so on. The retrieval problem is formulated as the problem of searching in a large pictorial database for images having features coinciding with or closely similar to those of a given example (training image).

Most of the recently developed CBIR techniques are based on colour and texture features. Popular colour features include a pixel-wise colour histogram [6], colour moments [7], and colour set vectors [8, 9]. Experiments in [7] have shown that the moment-based features perform sometimes better than other ones.

Texture features describing spatial signal interrelations are also widely used in image retrieval. Most of such features measure statistical properties of grey level co-occurrences in particular subsets of pixels, e.g., contrast, inverse difference moment, entropy and several other properties of a grey level co-occurrence matrix in [10, 11], statistics for a group of such matrices reflecting a characteristic structure of pairwise pixel interactions [12], or statistics of coefficients of the wavelet or Gabor image transforms [13, 14]. Today's CBIR exploits usually first three of the six perceptually meaningful textural features derived from the co-occurrence statistics in [15], namely, image coarseness, contrast, directionality, linelikeness, regularity, and roughness.

Combinations of particular features result sometimes in better performance [16, 17, 18]. Thus one might expect that some combined colour-texture features can in principle enhance the current CBIR techniques. This paper proposes and investigates an algorithm that retrieves a given training image and/or similar images from a large image collection by comparing characteristic subsets of normalised colour co-occurrence histograms (nCCHs) collected over small windows of a fixed size around each pixel. The characteristic subsets are selected using a generic Gibbs random field (GGRF) texture model proposed initially in [19, 20] for greyscale spatially homogeneous textures.

2 Retrieval Algorithm: Basic Steps

The proposed algorithm converts first the original RGB colour space into the HSV (Hue, Saturation, Value)

one and quantise this latter in order to reduce the data volume while preserving most of the training colours. Then the GGRF model is used for selecting the pixel neighbourhood which is most characteristic for the training image and for collecting the corresponding training nCCHs both over the entire image (as to describe it quantitatively) and in the moving windows of a fixed size centred at each pixel. These latter nCCHs are used to select a similarity threshold for detecting pixels that might belong to the desired texture. The threshold relates to the minimum similarity between the local and global training nCCHs. After selecting the candidate texture regions, their similarity to the training texture is measured by forming an empirical distribution of distances between the local (window-based) nCCHs for the image at hand and the global (entire-image-based) training nCCHs and comparing it to the like empirical training distribution.

Let both the training sample and all the images in an image database be already converted into the HSV colour space. Then the basic steps of the algorithm are as follows:

1. Collect the colour set by quantising the training sample.
2. Choose the most characteristic nCCHs for the training sample.
3. Find the empirical distribution of the distances between the training local and global nCCHs.
4. Select the maximum distance as the distance threshold.
5. For each image in the database,
   (a) Select the candidate colour pixels in the image using the above training colour set.
   (b) Find the candidate texture regions by thresholding the distances between the local nCCHs in the moving window around each candidate colour pixel and the global training nCCHs.
   (c) Calculate the distance between the empirical distribution of the nCCHs-based distances over the candidate texture region and the like empirical training distribution.
6. Retrieve the database images with the bottom-rank distances being below a certain threshold.

2.1 Colour space quantization

The RGB colour space is the most common format for digital images, while the HSV (Hue, Saturation, Value) model is more attractive in CBIR applications because of perceptually meaningful independent channels. To match human colour perception having more tolerance to saturation and value deviations, the quantised images should preserve more hue levels comparing to the other two channels. The quantisation helps also to keep a reasonable computing time.

Let $\mathbf{G} = [\mathbf{g}_i = [g_{i,\mathrm{h}}, g_{i,\mathrm{s}}, g_{i,\mathrm{v}}] : i = 1,\ldots,M; \mathbf{g}_i \in \mathbf{Q}_{\mathrm{hsv}}]$ where $\mathbf{Q}_{\mathrm{hsv}} = \{\mathbf{Q}_{\mathrm{h}}, \mathbf{Q}_{\mathrm{s}}, \mathbf{Q}_{\mathrm{v}}\}$ denote a digital colour image in the 3D HSV vector colour space. Here, $\mathbf{g}_i$ is a colour vector for the image position i being a shorthand notation of the 2D integer Cartesian coordinates $i = (x,y)$, and $\mathbf{Q}_{\mathrm{h}}, \mathbf{Q}_{\mathrm{s}}, \mathbf{Q}_{\mathrm{v}}$ are the finite integer sets of the colour component values: $\mathbf{Q}_{\mathrm{h}} = \{0,\ldots,Q_{\mathrm{h}}\}$, $\mathbf{Q}_{\mathrm{s}} = \{0,\ldots,Q_{\mathrm{s}}\}$, $\mathbf{Q}_{\mathrm{v}} = \{0,\ldots,Q_{\mathrm{v}}\}$. Due to a specific biconic form of the perceived HSV space, the quantised colours involve $Q_{\mathrm{h}}+1$ hue levels, Q_{s} saturation levels and Q_{v} value levels plus $Q_{\mathrm{v}}+1$ pure grey levels, so that the total number of colours after quantisation is $\tau = (Q_{\mathrm{h}}+1)\times Q_{\mathrm{s}}\times Q_{\mathrm{v}} + Q_{\mathrm{v}} + 1$. For example, if $Q_{\mathrm{h}} = 17, Q_{\mathrm{s}} = 3, Q_{\mathrm{v}} = 3$, then $\tau = 166$.

Let $\mathbf{G}^{tr} = [\mathbf{g}_j^{tr} = [g_{j,\mathrm{h}}^{tr}, g_{j,\mathrm{s}}^{tr}, g_{j,\mathrm{v}}^{tr}] : j = 1,\ldots,M^{tr}; \mathbf{g}_j^{tr} \in \mathbf{Q}_{\mathrm{hsv}}]$ denote the training image. Let a colour set Φ present all the colours contained in the training sample $\mathbf{G}^{tr}$. Then a candidate image $\mathbf{G}' = [\mathbf{g}_i' : i = 1,\ldots,M; \mathbf{g}_i' \in \Phi]$ can be obtained from the image $\mathbf{G}$ to be searched for as follows:

$\mathbf{g}_i' = \mathbf{g}_i$, if $\mathbf{g}_i \in \Phi$ or it is sufficiently close to that set;

$\mathbf{g}_i' = [0, 0, Q_{\mathrm{v}}]$, otherwise (the white background).

The candidate image $\mathbf{G}'$ that keeps all the colours similar to ones in the training sample is used at the next step. In order to reduce computing time, the training sample $\mathbf{G}^{tr}$ and all the candidate images $\mathbf{G}'$ use the index images with respect to the colour set Φ, so that the images themselves need not be changed.

Let the index set of Φ be $\mathbf{Q} = \{0,\ldots,Q\}$. Then the index images of the training sample $\mathbf{G}^{tr}$ and the candidate image $\mathbf{G}'$ will be $\mathbf{G}^{tr} = [g_j^{tr} : j = 1,\ldots,M^{tr}; g_j^{tr} \in \mathbf{Q}]$ and $\mathbf{G}' = [g_i' : i = 1,\ldots,M; g_i' \in \mathbf{Q}]$, respectively.

2.2 Pixel-wise Similarity Measure

Spatially homogeneous greyscale image textures can be modelled as samples of a generic Gibbs random field (GGRF) with multiple pairwise pixel interactions [19, 20]. Characteristic geometric structure of interactions and Gibbs potentials giving quantitative interaction strengths for a particular texture are analytically estimated from the training sample of the texture. The estimation yields a characteristic subset of pixel neighbours $\mathbf{A}$ specifying most "energetic" translation invariant families of interacting pixel pairs, or cliques of the neighbourhood

graph, $\mathbf{C}_a = \{(i, i+a) : i \in \{1,\ldots,M\}; i+a \in \{1,\ldots,M\}\}$; $a \in \mathbf{A}$. In our case, each such clique family presented in the GGRF model is extended onto a colour texture by the corresponding CCH acting as a sufficient statistic.

Let $\mathbf{F}_a(\mathbf{G}^{tr}) = [F_a(q,s|\mathbf{G}^{tr}) : q,s \in \mathbf{Q}]$ and $\mathbf{F}_{a,i}(\mathbf{G}) = [F_{a,i}(q,s|\mathbf{G}) : q,s \in \mathbf{Q}]$ denote the global nCCH for the clique family $\mathbf{C}_a$ over the training sample $\mathbf{G}^{tr}$ and the like local nCCH over the moving window $\widetilde{\mathbf{W}}$ around a position i in the image $\mathbf{G}$, respectively [21]. Experiments with different textures show that the symmetric χ^2-distance between these two CCHs:

$$D_{a,i}(\mathbf{F}_{a,i}(\mathbf{G}), \mathbf{F}_a(\mathbf{G}^{tr})) = \sum_{q,s \in \mathbf{Q}} \frac{\left(F_{a,i}(q,s|\mathbf{G}) - F_a(q,s|\mathbf{G}^{tr})\right)^2}{F_{a,i}(q,s|\mathbf{G}) + F_a(q,s|\mathbf{G}^{tr})} \quad (1)$$

is much less scattered over the training sample than, for instance, the pixel-wise Gibbs energies or conditional probabilities of signals. Therefore, for the $|\mathbf{A}|$ characteristic families, the pixel-wise similarity measure between $\mathbf{G}$ and $\mathbf{G}^{tr}$ can be defined as follows:

$$D_i(\mathbf{F}_i(\mathbf{G}), \mathbf{F}(\mathbf{G}^{tr})) = \frac{1}{|\mathbf{A}|} \sum_{a \in \mathbf{A}} D_{a,i}(\mathbf{F}_{a,i}(\mathbf{G}), \mathbf{F}_a(\mathbf{G}^{tr})) \,.$$

In fact, most of the CCHs for the different clique families in a homogeneous texture have very similar patterns. Therefore the similarity measure does not change significantly when the number of the characteristic families is increasing. In our experiments below only one most "energetic" family is used to calculate the distances.

The candidate texture regions in the image $\mathbf{G}'$ are detected pixel-by-pixel by thresholding all the distances $[D_i(\mathbf{F}_i(\mathbf{G}'), \mathbf{F}(\mathbf{G}^{tr})) : i = 1,\ldots,M]$ using the distance threshold $\xi = \max\{D_j(\mathbf{F}_j(\mathbf{G}^{tr}), \mathbf{F}(\mathbf{G}^{tr})) : j = 1,\ldots,M^{tr}\}$. The detected regions are of the similar texture type relating to the training sample with respect to the pixel-wise CCH-based similarity measure.

2.3 Region Similarity Measure

Let $\Omega(\mathbf{G}')$ be the candidate texture regions in the image $\mathbf{G}'$. Let $\mathbf{D}(\Omega(\mathbf{G}')) = [D_k(\mathbf{F}_k(\mathbf{G}'), \mathbf{F}(\mathbf{G}^{tr})) : k \in \Omega(\mathbf{G}')]$ be the distance distribution over the candidate texture regions. Let $\mathbf{D}(\mathbf{G}^{tr})) = [D_j(\mathbf{F}_j(\mathbf{G}^{tr}), \mathbf{F}(\mathbf{G}^{tr})) : j = 1,\ldots,M^{tr}]$ be the distance distribution over the training sample. The distance range in Eq.(1) is from 0 to 2 inclusively. We can estimate relative frequency distributions $\mathbf{D}(\mathbf{G}^{tr}))$ and $\mathbf{D}(\Omega(\mathbf{G}'))$ by quantising the distance range with a certain step Δ, $\mathbf{U} = \{0, \Delta, 2\Delta, \ldots, U\Delta = 2\}\}$. The symmetric χ^2-distance between these two distributions similar to Eq.(1) gives the region similarity measure ζ such that $0 \le \zeta \le U$. The desired images to be retrieved have the top-rank ζ–values after the database is sorted by the descending distance order, that is, the smaller the distance ζ, the higher the rank of an image and the more likely the image regions similar to the training sample. A threshold $\bar{\zeta}$ of the region similarity measure ζ should be chosen to reject images with the "inappropriate" candidate regions.

3 Experimental Results and Conclusions

The experiments below use 167 colour images (128×128) from the MIT Media Lab VisTex database. Other parameters of the algorithm are chosen as follows: the moving window 17×17, one ($|\mathbf{A}| = 1$) most energetic clique family per each training sample $\mathbf{G}^{tr}$ selected among 1830 possible variants with the inter-pixel x and y offsets up to 30, and the quantising distance step $\Delta = 0.0125$. The original VisTex texture patches (64×64) selected as the training samples are shown in Fig. 1.

Figures 2 – 9 illustrate results of texture retrieval using the proposed algorithm and the HSV quantisation with $Q_{\rm h} = 23, Q_{\rm s} = 3, Q_{\rm v} = 3$. Each example demonstrates the original VisTex texture patch selected as the training sample (a) and the four top-rank (in similarity) retrieved images with quantitatively similar texture regions (b)-(e). White regions in the images represent the rejected background having colours that differ much from the training ones. As one may expect, the images have the smallest region distance with respect to their own training patches.

Table 1 lists more results of colour texture retrieval. If the threshold $\bar{\zeta} = 1.75$, almost all the retrieved images are our expected ones except for the training samples "Fabric.4", "Metal.0", and "Tile.0".

The reason is that the retrieval accuracy depends on the number of colours for the training image in the colour set Φ specified by the HSV quantization. Generally, texture discrimination becomes reasonably good for the chosen image database if the number of individual colours in the set Φ is greater than 16. Results of the colour texture retrieval for training sample "Fabric.4" and "Stone.4" with different variants of the HSV quantisation are shown in Tables 2-3. This suggests that the set Φ has to be found by dynamically choosing the quantising levels for each training sample in order to obtain the precise retrieval. Our experiments show that $16 \le |\Phi| \le 32$ is a suitable range for most of the images in our database. Meanwhile, an inappropriate set Φ can result in some unexpected retrieved images. For example, for the training sample "Brick.0" shown in Fig. 3, the totally different texture "Misc.2" emerges among the images retrieved with $Q_{\rm h} = 23, Q_{\rm s} = 3, Q_{\rm v} = 3$, and $|\Phi| = 8$. However, when $|\Phi| = 17$ under the quantisa-

tion with $Q_h = 27, Q_s = 5, Q_v = 5$, the texture "Misc.2" is be excluded from the retrieved subset, and only the two textures "Brick.0" ($\zeta = 0.049$) and "Brick.1" ($\zeta = 1.29$) remain as the retrieved ones. These and other our experiments show that the proposed algorithm effectively retrieves images containing the desired textures.

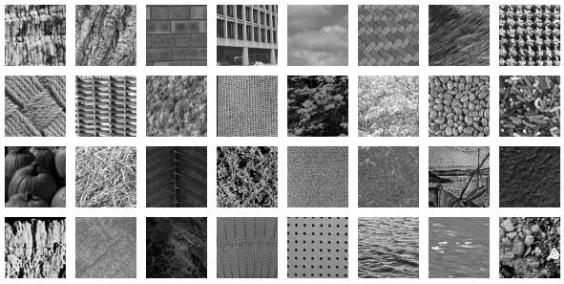

Figure 1: The original VisTex texture patches (64×64) selected as the training samples: Bark.0, Bark.12, Brick.0, Buildings.1, Clouds.0, Fabric.0, Fabric.4, Fabric.8; Fabric.11, Fabric.13, Fabric.15, Fabric.17, Flowers.0, Flowers.4, Food.0, Food.6; Food.10, Grass.1, Leaves.1, Leaves.12, Metal.0, Misc.2, Paintings.1.0, Sand.0; Stone.1, Stone.4, Terrain.0, Tile.0, Tile.7, Water.0, Water.2, WheresWaldo.1, respectively.

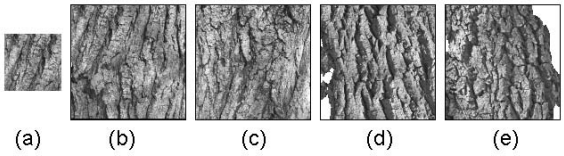

Figure 2: (a) Training sample Bark.12; (b)–(e) the retrieved four top-rank texture regions: Bark.12, $\zeta = 0.12$; Bark.11, $\zeta = 0.22$; Bark.10, $\zeta = 0.91$; Bark.9, $\zeta = 1.49$, respectively.

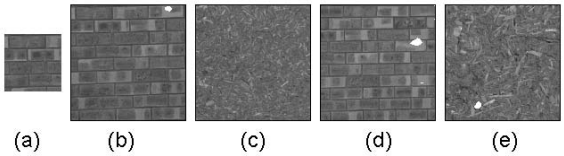

Figure 3: (a) Training sample Brick.0; (b)–(e) the retrieved four top-rank texture regions: Brick.0, $\zeta = 0.054$; Misc.2, $\zeta = 0.35$; Brick.1, $\zeta = 0.44$; Misc.3, $\zeta = 1.99$, respectively.

4 Acknowledgements

This work was supported in part by the Royal Society of New Zealand Marsden Fund under Grant 3600771/9143 (UOA 122).

References

[1] Y. Rui, T. S. Huang, and S. Chang. Image retrieval: current techniques, promising directions and open issues. *J. Visual Communication and Image Representation*, 10(4):39–62, 1999.

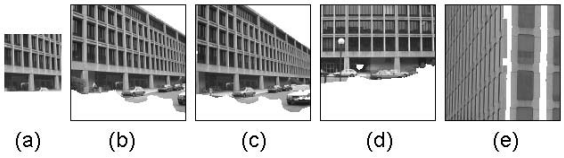

Figure 4: (a) Training sample: Buildings.1; (b)–(e) the retrieved four top-rank texture regions: Buildings.1, $\zeta = 0.084$; Buildings.2, $\zeta = 0.25$; Buildings.4, $\zeta = 0.84$; Buildings.10, $\zeta = 1.37$, respectively.

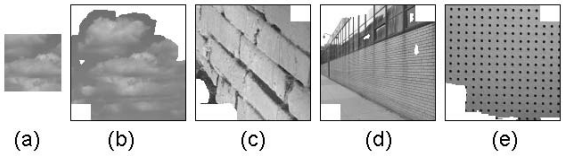

Figure 5: (a) Training sample: Clouds.0; (b)–(e) the retrieved four top-rank texture regions: Clouds.0, $\zeta = 1.36$; Brick.6, $\zeta = 1.82$; Buildings.6, $\zeta = 2.73$; Tile.7, $\zeta = 3.98$, respectively.

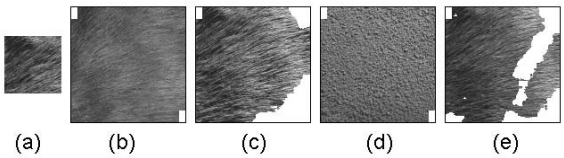

Figure 6: (a) Training sample: Fabric.4; (b)–(e) the retrieved four top-rank texture regions: Fabric.6, $\zeta = 4.98$; Fabric.4, $\zeta = 6.60$; Food.5, $\zeta = 13.24$; Fabric.5, $\zeta = 26.49$, respectively.

[2] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[3] S. K. Chang and A. Hsu. Image information systems: Where do we go from here? *IEEE Trans. Knowledge and Data Engineering*, 4(5):431–442, 1992.

[4] N. S. Chang and K. S. Fu. Picture query languages for pictorial database systems. *IEEE Computer Magazine*, 14(11):23–33, 1981.

[5] S. K. Chang and T. L. Kunii. Pictorial data-base systems. *IEEE Computer*, 14(11):13–21, 1981.

[6] M. J. Swain and D. H. Ballard. Color indexing. *Int. J. Computer Vision*, 7(1):11–32, 1991.

[7] M. Stricker and M. Orengo. Similarity of color images. In *Proc. SPIE Storage and Retrieval for Image and Video Databases*, volume 2420, pages 381–392, 1995.

[8] J. R. Smith and S. Chang. Single color extraction and image query. In *Proc. IEEE Int. Conf. Image Processing (ICIP'95)*, pages 528–531, 1995.

Table 1: Colour texture retrieval from the MIT VisTex database with $Q_{\mathrm{h}} = 23, Q_{\mathrm{s}} = 3, Q_{\mathrm{v}} = 3$.

Training Patch (64×64)	**Number of Colours**	**Three Top Rank Textures**					
		1		**2**		**3**	
		ζ	**Texture**	ζ	**Texture**	ζ	**Texture**
Bark.0	14	0.40	Bark.0	5.59	Bark.2	7.72	Bark.1
Bark.12	10	0.12	Bark.12	0.22	Bark.11	0.91	Bark.10
Brick.0	8	0.054	Brick.0	0.35	Misc.2	0.44	Brick.1
Buildings.1	10	0.084	Buildings.1	0.25	Buildings.2	0.84	Buildings.4
Clouds.0	4	1.36	Clouds.0	1.82	Brick.6	2.73	Buildings.6
Fabric.0	10	0.53	Fabric.0	7.44	Fabric.1	-	-
Fabric.4	9	4.98	Fabric.6	6.60	Fabric.4	13.24	Food.5
Fabric.8	12	0.095	Fabric.8	1.46	Fabric.9	1.48	Fabric.10
Fabric.11	10	0.14	Fabric.11	0.32	Fabric.12	10.98	Tile.9
Fabric.13	11	0.10	Fabric.13	0.21	Fabric.14	9.93	Fabric.4
Fabric.15	14	0.014	Fabric.15	0.031	Fabric.16	19.16	Paintings.41.1
Fabric.17	5	0.029	Fabric.17	1.17	Clouds.1	1.55	Brick.6
Flowers.0	25	0.012	Flowers.1	0.034	Flowers.0	5.12	Sand.2
Flowers.4	13	0.086	Flowers.4	0.23	Flowers.5	1.18	Flowers.7
Food.0	12	0.169	Food.0	11.51	Grass.0	11.97	Tile.9
Food.6	73	0.23	Food.6	0.31	Food.7	2.07	WheresWaldo.2
Food.10	14	0.35	Food.10	4.83	Food.11	19.68	Sand.4
Grass.1	13	0.57	Grass.1	2.13	Grass.2	9.65	Brick.7
Leaves.1	9	1.41	Leaves.1	5.88	Leaves.0	13.41	Bark.6
Leaves.12	13	1.47	Leaves.12	2.33	Leaves.13	13.77	Leaves.11
Metal.0	6	1.99	Metal.0	3.08	Metal.3	3.43	Metal.2
Misc.2	6	0.023	Misc.2	1.21	Misc.3	1.49	Stone.5
Paintings.1.0	49	0.54	Paintings.1.0	1.20	Paintings.31.1	1.81	Leaves.9
Sand.0	7	1.49	Sand.0	5.90	Sand.1	8.74	Sand.4
Stone.1	21	1.75	Stone.1	16.77	Brick.3	21.13	Sand.6
Stone.4	3	0.44	Water.5	0.44	Stone.4	0.79	Fabric.7
Terrain.0	11	0.31	Terrain.0	3.33	Terrain.9	3.51	Terrain.1
Tile.0	9	5.98	Tile.0	14.35	Metal.0	18.43	Metal.3
Tile.7	5	1.36	Tile.7	13.12	Clouds.6	16.14	Buildings.6
Water.0	6	0.11	Water.0	0.15	Water.7	0.19	Water.1
Water.2	10	0.47	Water.2	-	-	-	-
WheresWaldo.1	26	0.049	WheresWaldo.1	5.69	Food.4	6.49	Leaves.14

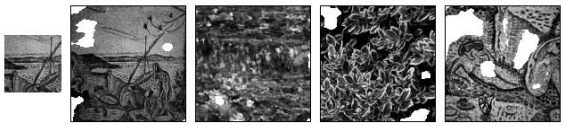

Figure 7: (a) Training sample: Paintings.1.0; (b)–(e) the retrieved four top-rank texture regions: Paintings.1.0, $\zeta = 0.54$; Paintings.31.1, $\zeta = 1.20$; Leaves.9, $\zeta = 1.81$; Paintings.1.1, $\zeta = 2.19$, respectively.

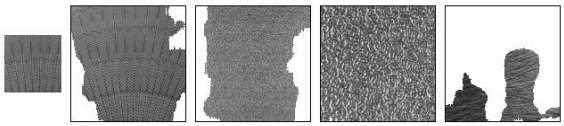

Figure 8: (a) Training sample: Tile.0; (b)–(e) the retrieved four top-rank texture regions: Tile.0, $\zeta = 5.98$; Metal.0, $\zeta = 14.35$; Metal.3, $\zeta = 18.43$; Fabric.5, $\zeta = 18.97$, respectively.

[9] J. R. Smith and S. Chang. Tools and techniques for color image retrieval. In *Proc. SPIE Storage and Retrieval for Image and Video Databases*, volume 2670, pages 426–437, 1996.

[10] C. C. Gotlieb and H. E. Kreyszig. Texture descriptors based on co-occurrence matrices. *Computer Vision, Graphics, and Image Processing*, 51(1):70–86, 1990.

[11] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE*

Table 2: Colour texture retrieval in VisTex database for training sample “Fabric.4” with different quantization in HSV colour space.

HSV Quantization	Number of Colours	Three Top Rank Textures					
		1		2		3	
		ζ	Texture	ζ	Texture	ζ	Texture
$Q_h = 23, Q_s = 3, Q_v = 3$	9	4.98	Fabric.6	6.60	Fabric.4	13.24	Food.5
$Q_h = 25, Q_s = 4, Q_v = 4$	14	1.19	Fabric.4	5.82	Fabric.5	8.99	Fabric.6
$Q_h = 27, Q_s = 5, Q_v = 5$	18	1.03	Fabric.4	8.13	Fabric.5	18.65	Food.5

Table 3: Colour texture retrieval in VisTex database for training sample “Stone.4” with different quantization in HSV colour space.

HSV Quantization	Number of Colours	Three Top Rank Textures					
		1		2		3	
		ζ	Texture	ζ	Texture	ζ	Texture
$Q_h = 23, Q_s = 3, Q_v = 3$	3	0.44	Water.5	0.44	Stone.4	0.79	Fabric.7
$Q_h = 25, Q_s = 4, Q_v = 4$	8	0.54	Stone.4	0.93	Fabric.18	1.51	Brick.3
$Q_h = 27, Q_s = 5, Q_v = 5$	10	0.63	Stone.4	3.88	Fabric.7	4.14	Fabric.18

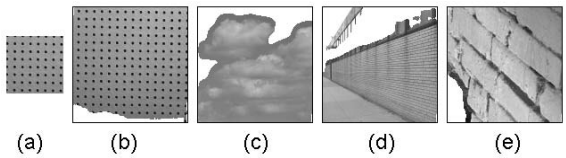

Figure 9: (a) Training sample: Tile.7; (b)–(e) the retrieved four top-rank texture regions: Tile.7, $\zeta = 1.36$; Clouds.0, $\zeta = 13.12$; Buildings.6, $\zeta = 16.14$; Brick.6, $\zeta = 16.40$, respectively.

Trans. Systems, Man and Cybernetics, 3(6):610–621, 1973.

[12] G. L. Gimel’farb and A. K. Jain. On retrieving textured images from an image data base. *Pattern Recognition*, 29(9):1461–1483, 1996.

[13] J. R. Smith and S. Chang. Transform features for texture classification and discrimination in large image databases. In *Proc. IEEE Int. Conf. Image Processing (ICIP’94)*, volume 3, pages 407–411, 1994.

[14] J. R. Smith and S. Chang. Automated binary texture feature sets for image retrieval. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, pages 2239–2242, 1996.

[15] H. Tamura, S. Mori, and T. Yamawaki. Texture features corresponding to visual perception. *IEEE Trans. Systems, Man and Cybernetics*, 8(6):460–473, 1978.

[16] M. H. Gross, R. Koch, L. Lippert, and A. Dreger. Multiscale image texture analysis in wavelet spaces. In *Proc. Int. Conf. Image Processing (ICIP’94)*, volume 3, pages 412–416, 1994.

[17] A. Kundu and J. Chen. Texture classification using qmf bank-based subband decomposition. *CVGIP: Graphical Models and Image Processing*, 54(5):369–384, 1992.

[18] K. S. Thyagarajan, T. Nguyen, and C. E. Persons. A maximum likelihood approach to texture classification using wavelet transform. In *Proc. IEEE Int. Conf. Image Processing (ICIP’94)*, volume 2, pages 640–644, 1994.

[19] G. L. Gimel’farb. Texture modeling by multiple pairwise pixel interactions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:1110–1114, 1996.

[20] G. L. Gimel’farb. *Image Textures and Gibbs Random Fields*. Kluwer Academic Publishers, Dordrecht, 1999.

[21] G. L. Gimel’farb and L. Yu. Separating a texture from an arbitrary background using pairwise grey level cooccurrences. In *Proc. 4th Int. Workshop Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 306–322, Portugal, Lisbon, July 2003. Springer, Berlin. Lecture Notes in Computer Science 2683.