Fast and Robust Image Identification

Paul Brasnett and Miroslaw Bober
Mitsubishi Electric ITE-VIL, Guildford, UK
{*paul.brasnett, miroslaw.bober*}@*vil.ite.mee.com*

Abstract

This paper presents an image identifier robust to common modifications. A multi-resolution Trace transform is introduced that constructs a set of 1D representations of an image. A binary identifier is extracted from each representation using a Fourier transform. Experimental evaluation of the algorithm and three state-of-the-art methods was carried out on a set of over 60,000 unique images with results demonstrating that the method outperforms the prior art methods in terms of detection, robustness and speed and achieves detection rate over 99% at a false-positive rate below 1 per million, with search speed exceeding 10 million image pairs per second on a desktop PC.

1. Introduction

A survey of prominent web sites show that Flickr has over 2 billion images, Photobucket has over 4 billion and Facebook has 1.7 billion. Consumers often have thousands of photographs in their personal collections and professionals will have significantly more. At this time there is a lack of robust tools supporting identification of visual multimedia content, for example to find the original version of a modified image. Such content identification tools are important to applications such as copyright protection, database de-duplication and management, content linking and content identification. Given the large size of datasets the tools must be very fast and deliver good detection at very low false-alarm rates.

Typically multimedia identification involves extracting an identifier that in some way captures the features. An image identifier must be robust to common image modifications. It should also be compact and it must allow extremely fast searching. Past work in the area of image identifiers can be classified into three approaches by their support region, i) local feature point based, ii) region based and iii) global.

Local feature methods offer a high level of robustness to geometric transformations [6] but generally have high complexity search. Region based approaches overcome some of the searching complexity issues of feature points but have poor performance particularly in the presence of rotation [8, 4]. Global methods have shown promise in terms of search complexity and robustness [5, 7, 2].

Local feature based methods are usually robust to geometrical transformations and cropping. However, brute force searching suffers from combinatorial explosion as the number of features increases. Methods do exist to improve performance but search costs are still relatively high.

The Trace transform is used in this work and develops from past work [2] by introducing the multi-resolution Trace transform which improves the results significantly. The contributions of this paper are i) introduction of the multi-resolution Trace transform, ii) mathematical justification of the identifier introduced in [2] and iii) experimental comparison using MPEG-7 test conditions with 3 alternative methods.

The Trace transform, its properties and how it can be used to extract invariant features is given in Section 2. The multi-resolution Trace transform is presented in Section 3. Extraction of the robust identifier from an image is described in Section 4. Section 5 presents the experimental results and then conclusions are in Section 6.

2 The Trace Transform

The Trace transform projects all lines over an image and applies functionals [1] over these lines. A further functional, known as the diametrical functional, is applied to the Trace transform to obtain a 1D function known as the circus function. An image identifier is developed using the trace and diametrical functionals.

[1]A functional is a real-valued function on a vector space V, usually of functions. In the case of the Trace transform the functionals are applied over the lines projected through the image $f(x, y)$.

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

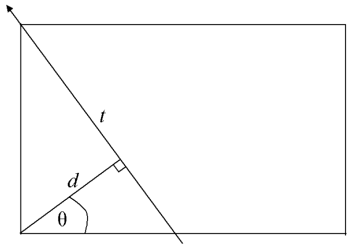
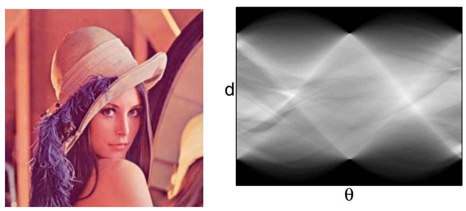


Figure 1. The Trace transform projects lines over the image. The lines are parameterised by the angle θ and distance d.



(a) Lena (b) Trace Transform (IF_1)

Figure 2. Shown here is Lena (a) and a Trace transform of Lena (b) using functional IF_1.

A line is parameterised in a co-ordinate system C_1 by (θ_1, d_1, t_1), see Fig. 1. Where θ_1 is the angle of the normal to the line, d_1 is the distance between the origin and line and t_1 is the distance along the line. The values of the image function along a particular line are

$$F_1(\theta_1, d_1, t_1) = F(C_1; \theta_1, d_1, t_1).$$

Now the Trace transform T applies some functional over the image function that results in the diametrical function

$$d(C_1; \phi_1, p_1) = T(F(C_1; \phi_1, p_1, t_1)).$$

An example of an image and its Trace transform is shown in Fig. 2. The diametrical functional D operates on the diametrical function to give the circus function

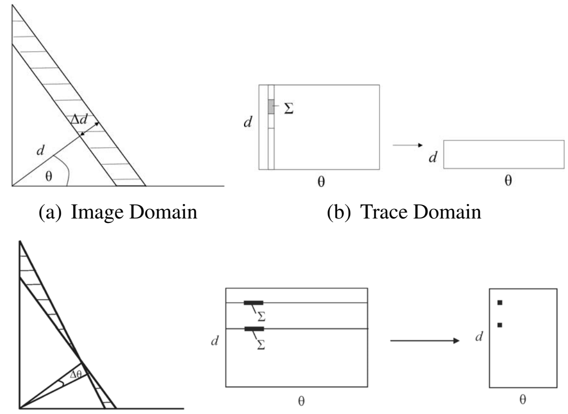
$$c(C_1; \phi_1) = D(T(F(C_1; \phi_1, p_1, t_1)). \qquad (1)$$

Invariant Functionals: A functional Ξ of a function $\xi(x)$ is invariant if

$$\Xi(\xi(x+b)) = \Xi(\xi(x)), \quad \forall b \in \mathbb{R}. \qquad (I_1)$$

Invariant functionals can have two further properties

$$\Xi(\xi(ax)) = \alpha(a)\Xi(\xi(x)), \quad \forall a > 0, \qquad (i_1)$$



(a) Image Domain (b) Trace Domain

(c) Image Domain (d) Trace Domain

Figure 3. Multi-resolution Trace transform.

$$\Xi(d\xi(x)) = \gamma(d)\Xi(\xi(x)), \quad \forall d > 0. \qquad (i_2)$$

It can be shown [3] that

$$\alpha(a) = a^{\kappa_\Xi}, \gamma(d) = d^{\lambda_\Xi},$$

where κ_Ξ and $\gamma(d) = d^{\lambda_\Xi}$ are characteristic properties of the functional Ξ. The following functionals obey properties I_1, i_1 and i_2

$$\begin{aligned} IF_1 &= \int \xi(t)\mathrm{d}t, \\ IF_3 &= \int |\xi(t)'|\,\mathrm{d}t, \\ IF_6 &= \max(\xi(t)). \end{aligned}$$

3 Multi-resolution Trace Transform

Multi-resolution representations are popular for their powerful ability to describe signals at varying levels of detail from coarse to fine. Here a multi-resolution Trace-transform is introduced that is quickly and efficiently generated from the original Trace transform.

A Trace transform $\boldsymbol{T}$ with a specific functional provides one representation of an image. From this one abstraction a multi-resolution representation of the image can be generated which captures information at different scales. The Trace transform multiresolution decomposition is performed by sub-sampling the original Trace transform of the image in either of its two dimensions, d or θ, or in both dimensions. The decomposition involves only a series of 1D operations, which can be carried out very efficiently with a small additional complexity. Decomposition in just one dimension is of order $O(N)$ where N is the magnitude of the dimension of the Trace transform and in both dimensions $O(N+M)$ where the magnitude of the dimensions are N and M. In practice it only adds 5% in extraction complexity and does not impact matching complexity, while greatly improving the performance.

The decomposition is performed in the *trace-domain* by sub-sampling the d-parameter by integrating over intervals along the columns, as in Figure 3(b). This corresponds to projecting strips of width d over the image during the Trace transform, as shown in Figure 3(a).

Sub-sampling also take place by integrating over intervals in the θ parameter(Fig. 3c), that is along the rows. As is shown in Fig. 3d this is approximately equivalent to integrating over double-cones in the image with opening-angle θ during the trace transform.

4 Robust Identifier

An image $f(x, y)$ can be viewed from two different co-ordinate systems C_1 and C_2. The coordinate system, C_2, is obtained from the C_1 by a rotation of angle $-\phi$, scaling the axis by parameter υ and by translating with the vector $(-s_0 \cos\Psi_0, -s_0 \sin\Psi_0)$. The image $f_2(\tilde{x}, \tilde{y})$ viewed from C_2 can be seen as the image $f_1(x, y)$ having undergone rotation by ϕ, scaling by υ^{-1} and shifting by $(s_0 \cos\Psi_0, s_0 \sin\Psi_0)$. Under these linear transformations a line in f_1 will still be a line in f_2; the transformations are line preserving. The parameters of an image line in co-ordinate system C_1 in terms of the parameters of the line in C_2 are

$$\begin{aligned}\theta_1 &= \theta_2 - \phi,\\ d_1 &= \upsilon\left[d_2 - s_0\cos(\Psi_0 - \theta_2)\right],\\ t_1 &= \upsilon\left[t_2 - s_0\sin(\Psi_0 - \theta_2)\right].\end{aligned}$$

From (1) it can be seen that the relationship of the circus function of an image in co-ordinate system C_2 to the image in coordinate system C_1 is given as

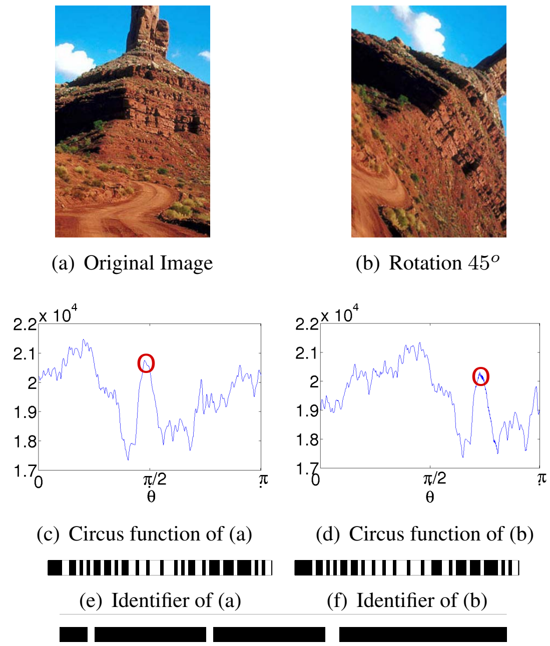
$$\begin{aligned}c(C_2;\phi_2) = D(T(F_1(\phi_1{-}\theta, \upsilon\left[p_1 - s_0\cos(\Psi_0 - \phi_1)\right],\\ \upsilon\left[t_1 - s_0\sin(\Psi_0 - \phi_1)\right]))).\end{aligned} \quad (2)$$

The Trace functional T is chosen to obey (I_1) and (i_1)

$$\begin{aligned}c(C_2;\phi_2) = D(\alpha_T(\upsilon)T(F_1(\phi_1 - \theta,\\ \upsilon\left[p_1 - s_0\cos(\Psi_0 - \phi_1)\right], t_1))).\end{aligned} \quad (3)$$

Furthermore, the diametrical functional D can be chosen to obey (I_1), (i_1) and (i_2) such that

$$\begin{aligned}c(C_2;\phi_2) = \gamma_D(\alpha_T(\upsilon))D(T(F_1(\phi_1 - \theta,\\ \upsilon\left[p_1 - s_0\cos(\Psi_0 - \phi_1)\right], t_1))),\end{aligned}$$



(a) Original Image (b) Rotation 45^o

(c) Circus function of (a) (d) Circus function of (b)

(e) Identifier of (a) (f) Identifier of (b)

(g) Identifier Difference

Figure 4. Circus function (c) and identifier (e) of image (a) and for image rotated by 45^0 (b,d,f). The circus function is shifted to the right by 45^0 $(\pi/4)$. The difference between identifiers is shown in (g).

$$\begin{aligned}c(C_2;\phi_2) = \gamma_D(\alpha_T(\upsilon))\alpha_D(\upsilon)D(T(F_1(\phi_1{-}\theta, p_1, t_1))),\\ c(C_2;\phi_2) = \kappa D(T(F_1(\phi_1 - \theta, p_1, t_1))),\end{aligned} \quad (4)$$

where $\kappa = \gamma_D(\alpha_T(\upsilon))\alpha_D(\upsilon)$. From (4) it can be seen that the 1D circus function in C_2 is a scaled and shifted version of the circus function in C_1. Example circus function are shown in Figure 4.

Identifier Extraction: Taking the Fourier transform of (4) gives

$$F(\Phi) = \mathcal{F}\left[\kappa D(T(F_1(\phi_1 - \theta, p_1, t_1)))\right],$$

then exploiting the linearity identity and translation property of the Fourier Transform gives

$$F(\Phi) = \kappa \exp^{-j\theta\Phi} \mathcal{F}\left[D(T(F_1(\phi_1, p_1, t_1)))\right].$$

Taking the magnitude of $F(\Phi)$ gives

$$|F(\Phi)| = |\kappa\mathcal{F}\left[D(T(F_1(\phi_1, p_1, t_1)))\right]|. \quad (5)$$

Using the properties of the circus function (4) and the magnitude of the Fourier transform an identifier can be extracted from an image. An algorithm to extract the robust binary identifier is given in Table 1. The identifier is robust under similarity transform, that is scaling, rotation and translation.

Table 1. Identifier Extraction Algorithm

1. Take the Trace transform of the image $f(x, y)$ using an appropriately chosen functional.
2. Obtain the circus function by applying a diametrical functional to the columns from the previous step.
3. Take the magnitude of the Fourier transform of the circus function $|F(\omega)|$.
4. A binary string is obtained by taking the difference between neighbouring coefficients

$$c(\omega) = |F(\omega)| - |F(\omega + 1)|, \quad (6)$$

   and then applying a threshold such that

$$b_\omega = \begin{cases} 0, & c(\omega) < 0 \\ 1, & c(\omega) \geq 0 \end{cases}. \quad (7)$$

5. The first bit b_0 corresponding to the DC component is discarded and the identifier is made up of the subsequent N bits,

$$B = [b_1, \ldots, b_N]. \quad (8)$$

6. For each diametrical functional perform steps (2)-(5).
7. Concatenate each of the identifiers to obtain the complete identifier.

Multiresolution Identifier: The multi-resolution Trace transform provides more identifiers. A 1D decomposition over the distance (d) parameter is performed. The extraction process shown in Table 1, Steps 2-5, are used for each level of the multi-resolution Trace transform. Significant performance improvements are obtained by extracting multiple identifiers from each image. Firstly different identifiers are extracted by making different choices for the diametrical functionals in Steps 1 and 2 of Table 1.
Distance: The distance between identifiers B_1 and B_2, of length N, is the normalised Hamming distance

$$H(B_1, B_2) = \frac{1}{N} \sum_N B_1 \oplus B_2, \quad (9)$$

where $\oplus$ is the exclusive OR (XOR) operator.

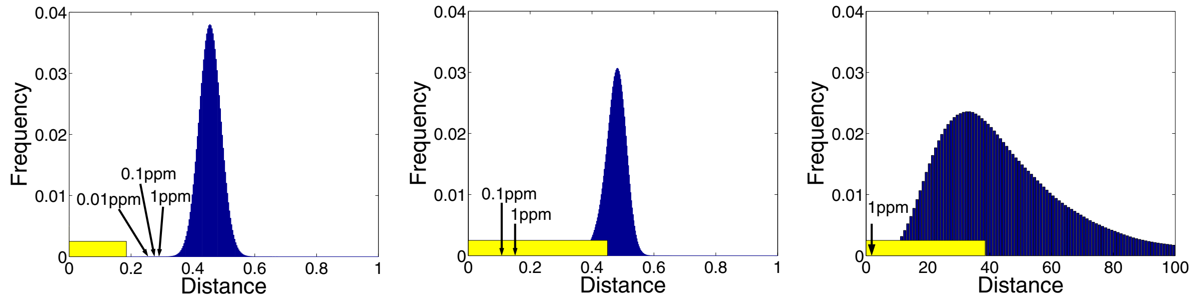
5 Results

The method developed in this paper (MRT, for Multi-Resolution Trace) is compared with three alternative methods. The first is a related technique based on the Radon transform [7]. The second is a region based approach [4] that uses pseudo-random sampling followed by a two stage SVD. The third method is a single-resolution Trace transform [2].

Figure 5. Six examples of the 60,000 image dataset used to evaluate the independence of the methods.

For the MRT method evaluated here, 12 basic identifiers are extracted, each of 48 bits giving a total of 576 bits. The Trace functional chosen is IF_1, and a multi-resolution Trace transform is formed by decomposing the original matrix in the d-parameter by factors of 8, 16, 32, 64 and 128. The identifiers are extracted by applying diametricals IF_3 and IF_6 to each of the six levels. The multi-resolution components are obtained directly in the trace domain and therefore add less then 5% extraction complexity. A bit selection procedure reduces the number of bits from 576 to 296 [2].

Experiments follow the procedure in MPEG-7 for visual identifiers [1]. It involves two tests, i) independence of over 60,000 independent images (1.8 billion pairs) and ii) robustness to typical image modifications at a false acceptance rate (FAR) of 1 part per million (1ppm). The dataset is made up of a variety of types including photographs, artworks and graphics.
Complexity: Extraction is dominated by the Trace transform which is of order $O(NMR)$ for an $N \times M$ image sampled at R angles. The current implementation takes 0.12s per image on a P4 3GHz PC, the other methods are comparable. For searching the Trace, MRT and Radon methods use the Hamming distance which, for an N bit identifier, requires N-XORs and a summation per comparison. The SVD method uses the L2-norm so for an identifier of length N requires N floating point subtraction, N multiplies, a summation and a square root. Clearly the SVD method will be significantly slower than the other methods when searching.
Independence: Histograms of the normalised Hamming distances between the independent images for three methods are presented in Figure 6. The thresholds corresponding to various FARs are shown.
Robustness: The independence threshold corresponding to a 1ppm FAR is found. The detection rates at this point for the different methods are presented in Table 2. It can be seen that the MRT method proposed



(a) MRT (b) Radon (c) SVD

Figure 6. Histograms of the distance between identifiers from 1.8 billion independent image pairs, FARs are shown with arrows. The rectangle highlights the distance between 90% of the modified images.

outperforms all of the prior art methods achieving the mean performance of 99.28%. The second best method, the single resolution trace transform reaches a detection rate of 92.40%.

Table 2. Detection rate (%), tested on 10,000 images, at false acceptance rate of 1 part per million.

	Trace	*MRT*	Radon	SVD
Blur 3	98.94	**100.00**	99.09	5.62
Blur 5	95.44	**100.00**	97.67	3.39
Bright $+5\%$	99.19	**100.00**	100.00	34.78
Bright $+10\%$	95.74	**99.80**	99.85	19.73
Bright $+20\%$	82.61	**97.87**	92.28	9.90
Flip (LR)	92.52	**97.01**	0.91	0.41
JPEG 95%	100.00	**100.00**	100.00	100.00
JPEG 80%	99.52	**100.00**	100.00	58.38
JPEG 65%	98.73	**100.00**	100.00	39.28
GIF	90.44	**99.75**	99.26	22.08
Greyscale	99.52	**100.00**	99.97	100.00
Noise $\sigma = 8$	92.09	**99.97**	100.00	48.89
Rotate 10^o	87.17	**98.00**	0.68	0.51
Rotate 25^o	87.63	**98.02**	0.51	0.33
Rotate 45^o	87.25	**97.64**	0.41	0.26
Scale 50%	77.99	**99.52**	99.52	12.15
Scale 70%	89.73	**99.80**	99.85	16.67
Scale 90%	88.67	**99.67**	100.00	23.89
Mean	92.40	**99.28**	77.22	27.57

6 Conclusions

This paper introduces a multi-resolution extension to the Trace transform for the first time. It allows efficient extraction of multiple identifiers representing an image at different scales. All identifiers are combined to give a single compact and robust binary image identifier. Experimental evaluation has been carried out on using MPEG-7 test scenarios, a comparison showed that the developed algorithm improves performance against three state-of-the-art methods. The developed method achieves detection rate above 99% at $1ppm$ false alarm rate and enables very fast matching of over 10 million image identifiers per second. We also expect the novel multiresolution Trace transform proposed here to be applicable other image processing applications.

Future work will investigate efficient extraction of additional robust information an image without compromising on search speeds. This work offers very low detection rates suitable for datasets with 10s or 100s of millions of images, even lower rates will be needed to be of use on the very large datasets with billions of images that exist on the web.

References

[1] M. Bober and R. Oami. Description of mpeg-7 visual core experiments. Technical Report N8971, ISO/IEC JTC1/SC29/WG11, Apr. 2007.

[2] P. Brasnett and M. Bober. A robust visual identifier using the trace transform. In *VIE 2007*, Jul. 2007.

[3] A. Kadyrov and M. Petrou. The trace transform and its applications. *IEEE Trans. on PAMI*, 23(8), Aug. 2001.

[4] S. Kozat, R. Venkatesan, and M. Mihçak. Robust perceptual image hashing via matrix invariants. In *IEEE Int. Conf. Image Proc.*, pages 3443–3446, Oct. 2004.

[5] C. D. Roover, C. D. Vleeschouwer, F. Lefèbvre, and B. Macq. Robust video hashing based on radial projections of key frames. *IEEE Trans. Sig. Proc.*, 53(10):4020–4037, Oct. 2005.

[6] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Trans. on PAMI*, 5(19), 1997.

[7] J. S. Seo, J. Haitsma, T. Kalker, and C. D. Yoo. A robust image fingerprinting system using the Radon transform. *Signal Processing: Image Comm.*, 19:325–339, 2004.

[8] R. Venkatesan, S.-M. Koon, M. Jakubowski, and P. Moulin. Robust image hashing. In *IEEE Int. Conf. Image Proc.*, pages 664–666, Sep. 2000.