A Probabilistic Model for Random Binary Image Mapping

ADNAN A. Y. MUSTAFA
Department of Mechanical Engineering
Kuwait University
P. O. Box 5969 - Safat – 13060
KUWAIT
adnan.nustafa@ku.edu.kw

Abstract: - Many probabilistic models have been developed for numerous problems in robot and computer vision such as for image segmentation, road extraction, and object tracking. In this paper we present a probabilistic model for the random pixel mapping of binary images. The model predicts the probability of detecting dissimilarity between dissimilar binary images as a function of the number of random mappings and the amount of similarity. The model shows that detecting dissimilarity can be accomplished quickly by random pixel mapping, without the need to process the entire image. Test results on real images are presented that show the accuracy of the model.

Key-Words: - Probabilistic model, binary images, pixel mapping, image matching, image mapping, dissimilarity detection, quick matching, big images.

1 Introduction

Image (and sub-image) matching rises frequently in the field of robot and computer vision under many topics such as, image registration [1] [2], template matching [3] [4], image retrieval [5] [6], image classification [7] [8], motion tracking [9], motion estimation [10], defect detection [11], …, etc. These methods are either feature-based methods [12] [13], that rely on extracting image features [1] [14] and then matching them, or area-based methods (aka intensity methods) [15] [16] [10] [11], based on comparing image intensity values.

Binary images have only two intensity levels, and thus a limited amount of scene detail is present in the image. As a result, binary image matching is usually accomplished by calculating the correlation between the images [15] [4] or simply by subtracting the two images [16] [17]. These methods, as well as all methods presented in the literature require some type of similarity criteria to be evaluated over the entire images being matched. Hence, these methods are image-size dependent; as image size increases, so does processing time. With modern day applications demanding higher image resolution resulting in higher image sizes (> 50 Mega-pixel images) in many fields such as robotics, industrial applications and medical surgery, traditional image-size based methods, such as correlation, can become quite slow in processing such huge images, even with the speed of today's computers. The dependency of matching techniques on image size is a serious handicap to these techniques.

These methods, correlation and image subtraction, as well as the majority –if not all– methods cited in the literature, perform image matching by scanning images line by line, top to bottom, processing every pixel value from one image with its corresponding value at the other image. However, if we look at how the human brain performs the task of matching between two pictures, we find that it takes a more efficient approach which is completely different; search for differences is performed by glancing at the two pictures and identifying regions of dissimilarity. The region selection process is performed in a random manner and once a region of dissimilarity is detected, then that region becomes the *area of focus*, where additional detailed comparison is performed. This process of random point selection is the basis upon which our technique is based upon. The traditional "top-down scan the entire image" approach is completely abandoned, and hence valuable time is conserved. We show that our approach of dissimilarity detection by random pixel selection and comparison is an approach which can detect dissimilarity very quickly. The proof is given by a probabilistic model that shows the quickness of detecting dissimilarity by random mapping. The model is called a *Probabilistic Matching Model for Binary Images* (*PMMBI*). The model predicts the probability of detecting dissimilarity between any pair of binary images as a function of the number of random pixel mappings and the level of similarity between them. By randomly mapping image pixels, the model reveals the following unique and important advantages,

1. Dissimilar images can be detected quickly without the need to process the entire matched images. Even images that are highly similar and near-duplicate, but not exactly similar, can be amazingly detected by only comparing a minute fraction of the total size of the images.

2. Detecting dissimilarity is image size invariant; the size of the image has absolutely no effect on the dissimilarity detection process nor its quickness. Detecting dissimilarity among images of size, say 500 gega-pixels is detected just as fast as detecting it among 10 kilo-pixels images.

3. Dissimilarity detection quickness between two images can be used to estimate image similarity to a good degree without the need to process the entire image. Hence, with the aid of the model, matching can be performed magnitudes faster than employing traditional matching techniques that require comparing the entire images.

Fig. 1 shows a comparison of the dissimilarity detection performance between *PMMBI* based methods and image-size based methods. Image-size based methods refer to traditional methods with the "top-down scan the entire image" approach. The top graph shows a plot of detection time vs. image size. The plot reveals that detection time increases linearly as a function of image size for image-size based methods, whereas detection time for *PMMBI* based methods is constant and does not depend on image size. The bottom graph shows a plot of detection time vs. image similarity. Detection time is constant for image-size based methods and does not increase with similarity, whereas detection time for *PMMBI* based methods increases as the amount of similarity between the images increases. However, detection time is very small for low similarity and does not become significant except at very high similarity. But even when images have high similarity, the detection time for *PMMBI* based methods is still smaller than that for image-size based methods.

This paper is organized as follows: section 2 points out related literature, while section 3 reviews related work necessary for the understanding of the model presented. Section 4 presents the main theme of this paper and presents the development of the probability *PMMBI* model. We discuss the model and show how it can be used to detect dissimilarity quickly without the need to process the entire image, but rather by mappings a few randomly
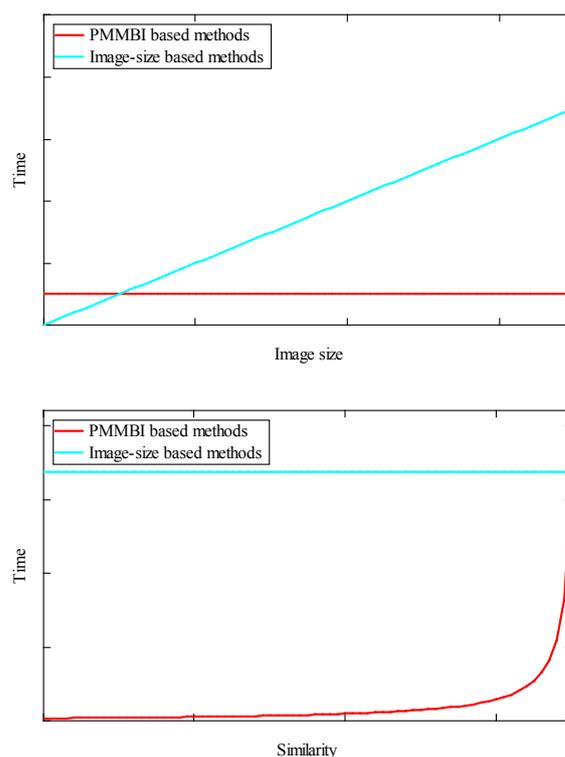

Fig. 1. Dissimilarity detection performance between *PMMBI* based methods and Image size based methods; top plot: time vs. image size. Bottom plot: time vs. similarity.

selected pixels. Section 5 presents results of detecting dissimilarity between images by random mapping that show the accuracy of the model. Section 6 presents our conclusion and where our future research is headed.

2 Related Literature

Image correlation [15] is a w ell established method for image matching and has been applied to many tasks [9] [10] [11] [18]. Current research on image correlation has focused on improving correlation calculations using a wide variety of techniques; such as using Fourier coefficients [19], pre-computed tables [4], updating computations at each window location [20], constructing basis functions [9], using Haar-like binary features [21], as well as numerous other techniques [22] [23] [24] [10] [25] [26] [27].

Image subtraction, usually computed as the sum of the absolute difference between two images [16] is also a common and popular approach for matching. It is the primary operation used for motion detection. Numerous techniques have also been developed to reduce the computations of the

matching problem [18] [28] [29]. Hardware implementations have also been proposed to speed up calculations [30] [31] [32] [33].

Mutual information is a third popular approach for matching that has seen growing interest over the last 20 years. It has been a very popular approach in the medical field as a solution to the image registration problem [1] [34] [35] [36].

Other area-based methods have also been developed based on a variety of principles; minimizing image intensity co-occurrences [37], using modified Hausdorff distances and local-dissimilarity maps [38], mathematical morphology [39], as well as other innovative similarity distances [40]. However, all of these methods are image size dependent and require that the entire images be processed for matching.

3 Related Work

In this section, we present a summary of some earlier developed concepts that are necessary for the understanding of the work presented in this paper. We present: 1) the definition of similarity between binary images, and how images are categorized based on it; 2) the γ similarity distance that is used in our work as an index for image similarity; 3) how dissimilarity detection is measured.

3.1 Similar and Dissimilar Binary Images

The closeness between two binary images is based on a pixel-to-pixel comparison between the binary images. Image closeness is categorized as either *similar* or *dissimilar* images [41] as follows:

- *Similar Images* (*S*): For images to be similar, the two images must be the same. They are of two types; either *exact* or *inverse*:
  - *Exact Images* (*E*): The two images are exactly the same; they have the same intensity values at all corresponding pixels.
  - *Inverse Images* (*I*): The two images are the inverse of each other, as they have the compliment intensity values at all corresponding pixels.

- *Dissimilar Images* (*R*): The two images are different and this can only be true if they are not similar; i.e. neither exact nor inverse. *Dissimilar images* are of two types:
  - *Distinct-dissimilar Images* (*D*): The two images are ideally different (as measured by γ –see below).

- *Quasi-dissimilar Images* (*Q*): The two images have concurrences between them at some pixels, but not all pixels. These images are also referred to as *Quasi-similar images*.

This categorization of binary images is the basis on which the probabilistic model discussed in this paper is based upon.

3.2 The Gamma Binary Similarity Distance

The *Gamma binary similarity measure* (γ) measures the amount of similarity and concurrence between two binary images [42] [43]. Formally stated: given two images **u** and **v**, γ is defined as,

$$\gamma(\mathbf{u},\mathbf{v}) = |1 - 2P_o((Z = \mathbf{u} \oplus \mathbf{v}) = z)|, \quad z \in \{0,1\} \quad (1)$$

where $\oplus$ is the *exclusive-or* operation and $P_o()$ denotes the probability mass function of the image intensities (i,e. the normal image histogram). As a result, $0 \leq \gamma \leq 1$, and hence values of γ correspond to,

- γ = 0 for *distinct-dissimilar* images
- 0 < γ < 1 for *quasi-dissimilar* images
- γ = 1 for *similar* images

In practice, image pairs with γ < 0.01 are assumed to be γ ≈ 0, and thus are considered to be *distinct-dissimilar* image pairs. Furthermore, image pairs with γ > 0.99 are termed as near-duplicate images, while images with γ > 0.999 are termed as near-similar images.

3.3 Measuring Mapping Performance

The *Mapping Detection Number* (*MDN*) is defined as the number of mappings required to detect a pair of images as being dissimilar. Furthermore, MDN_{DC} notation is used to denote *MDN* at a specific detection confidence (*DC*) value. For example, $MDN_{0.90} = 5$ implies that 5 mappings are sufficient to detect dissimilarity with 90% confidence.

4 The Probabilistic Matching Model for Binary Images

A variety of probabilistic models have been developed for numerous problems related to robot and computer vision; image segmentation [44], road extraction [45], vehicle detection [46], object tracking [47], image registration [48], image fusion [49], image comparison [50], image categorization [51], image retrieval [52], image tagging [53], and many more. In [41], a probabilistic model that predicts the probability of detecting dissimilarity between *distinct-dissimilar* binary images, called the

Probabilistic Matching Model (*PMM*) was presented. It showed that detecting dissimilarity between *distinct-dissimilar* binary images can be performed quickly by randomly selecting a few corresponding pixels between the matched images and comparing their values. More importantly, it showed that there is no need to compare entire images to detect dissimilarity. The model states that the probability of detecting dissimilarity, Pr(), between any pair of binary *distinct-dissimilar* images by the pth random mapping is given by the following formula,

$$\Pr(p) = \Pr(D, p) = 1 - \left(\frac{1}{2}\right)^{p-1}, \quad p = 2, 3, \ldots \quad (2)$$

It can be seen that Pr(*p*) approaches unity quickly after only a few mappings (*p*), e.g., Pr(5) = 0.938, Pr(8) = 0.992.

In this section, we introduce the Probabilistic Matching Model for Binary Images (*PMMBI*), which is a generalization of *PMM* that is applicable to any pair of binary images, not just distinct-dissimilar image pairs. Similar to *PMM*, *PMMBI* shows that selecting a few pixels randomly and mapping them between two images is sufficient to detect dissimilarity between them as long as the images are not near similar. In the latter case, more points need to be selected. The development of *PMMBI* is presented next.

4.1 Binary Pixel Mapping

Let **u** and **v** be two binary images, such that $u \in \mathbf{u}$ and $v \in \mathbf{v}$. *Binary Pixel mapping* (P_1) between two images refers to how a pixel value in the first image maps to the corresponding pixel value in the second image [54],

$$P_1 = \{u \rightarrow v \mid \forall u, v \in \{0,1\}\} \quad (3)$$

The '→' symbol is used to denote pixel mapping. Hence, $u \rightarrow v$ implies pixel value *u* in the first image maps to pixel value *v* in the second image. Since pixel values are either 0 or 1 for binary images, there are four possible pixel mappings between any pair of binary images,

$$P_1 = \{0 \rightarrow 0, 0 \rightarrow 1, 1 \rightarrow 0, 1 \rightarrow 1\} \quad (4)$$

Labels *A − D* are used as shorthand for these four mappings, respectively. Thus,

$$P_1 = \{A, B, C, D\} \quad (5)$$

Let *s* and *d* denote the set consisting of similar and different mapping values, respectively. Hence,

$$s = \{A, D\} \quad (6)$$
$$d = \{B, C\} \quad (7)$$

4.2 Binary Image Mapping

Assume that pixel locations are randomly selected and their intensity values are mapped. Let *k* be the probability of event *d* occurring at any given mapping; as a result the probability of *s* occurring is $(1 − k)$. On the first mapping two possible states are possible; *d* or *s*, as shown in **Fig. 2**; the probability of occurrence of *d* is *k* and the probability of *s* occurring is $(1 − k)$. On the second mapping, four cases are possible: *dd*, *ds*, *sd* and *ss*; their probabilities are k^2, $k(1 − k)$, $k(1 − k)$ and $(1 − k)^2$, respectively. On the third mapping there are 8 cases as shown in the figure. It can be seen that the probability distribution of *d* is a binomial distribution [55],

$$\varphi(X = x, p, k) = \binom{p}{x} k^x (1-k)^{p-x}, \quad x = 0, 1 \ldots p \quad (8)$$

where *X* is a random variable denoting the number of times *d* occurs on the pth mapping and φ is the probability of *d* occurring *x* times on the pth mapping. We refer to cases when both events occur (e.g. *ds*, *sddddds*) as mixed events. Since we are interested in the probability of occurrence of dissimilar images, then,

$$\Pr(k, p) = \varphi(0 < X < p, p, k)$$
$$= \sum_{x=1}^{p-1} \binom{p}{x} k^x (1-k)^{p-x} \quad (9)$$

represents the probability of occurrence of mixed events. This can be rewritten as,

$$\Pr(k, p) = 1 - \big(\varphi(X = 0, p, k) + \varphi(X = p, p, k)\big) \quad (10)$$

which simplifies to,

$$\Pr(k, p) = 1 - \big((1-k)^p + k^p\big) \quad (11)$$

4.3 The Probability of Occurrence of Detecting Dissimilar Binary Images

Let $\kappa^{(0)}(\mathbf{u}, \mathbf{v})$ be the probability that the intensity values of two images are the same at any point,

$$\kappa^{(0)}(\mathbf{u}, \mathbf{v}) \equiv p((\mathbf{z} = (\mathbf{u} \oplus \mathbf{v})) = 0) \quad (12)$$

where $p(\mathbf{z})$ is the probability mass function of the resulting image $\mathbf{z} = \mathbf{u} \oplus \mathbf{v}$. Similarly, define $\kappa^{(1)}(\mathbf{u}, \mathbf{v})$ as the probability that the intensity of two images have the compliment value at any point,

$$\kappa^{(1)}(\mathbf{u}, \mathbf{v}) \equiv p((\mathbf{z} = (\mathbf{u} \oplus \mathbf{v})) = 1) \quad (13)$$

Then, let κ be the image concurrence between two images defined as,

$$\kappa = \max(\kappa^{(0)}, \kappa^{(1)}) \quad (14)$$
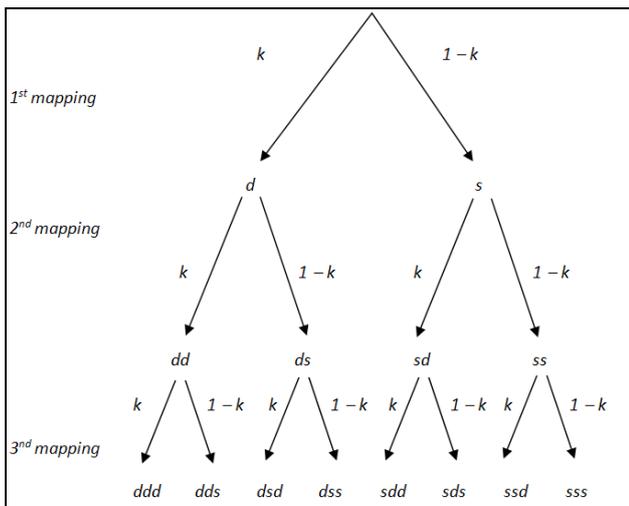
Fig. 2. Binary Image Mapping results up to 3 mappings.

i.e., κ is the probability that the intensity values of two images at any point are in agreement or disagreement, whichever is larger. The relation between γ and κ is given by [43],

$$\kappa(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\big(\gamma(\mathbf{u}, \mathbf{v}) + 1\big) \qquad (15)$$

But by definition, it can be seen that,

$$\kappa = k \qquad (16)$$

Substituting this relation in (15) and rearranging the equation produces,

$$k = (\gamma + 1) / 2 \qquad (17)$$

Finally, substituting this relation back into (11) and simplifying produces,

$$\Pr(\gamma, p) = 1 - \left(\frac{1}{2}\right)^p \big((1 + \gamma)^p + (1 - \gamma)^p\big) \qquad (18)$$

or

$$\Pr(\gamma, p) = 1 - \left(\frac{1}{2}(1 + \gamma)\right)^p \left(1 + \left(\frac{1 - \gamma}{1 + \gamma}\right)^p\right) \qquad (19)$$

where $0 \leq \gamma \leq 1$, $p = 2, 3, \ldots$. It is important to note that p is a discrete variable and γ is a continuous variable. This equation predicts the probability of detecting dissimilarity between any pair of binary images as a function of the number of mappings mapped thus far (p) and the amount of similarity (γ) between the images. This equation also implies that the probability of detecting dissimilarity is not a function of image size; dissimilarity detection is performed at the same speed (i.e. number of mappings) whether the image is in hundreds of

gigabytes or in a few kilobytes. Since the probability of detecting dissimilarity function, Pr(), is a measure of the confidence in detecting dissimilarity, it is also referred to as the *Detection Confidence* (*DC*).

We see that for distinct-dissimilar images (*D*), $\gamma = 0$ and thus (18) degenerates to,

$$\Pr(D, p) = \Pr(\gamma = 0, p) = 1 - 2\left(\frac{1}{2}\right)^p \qquad (20)$$

which agrees with the expression of *PMM* appearing in (2). At the other extreme, for similar images (*S*) which have $\gamma = 1$, (18) degenerates to,

$$\Pr(S, p) = \Pr(\gamma = 1, p) = 0 \qquad (21)$$

i.e., when images are similar, then there is no possibility of detecting dissimilarity between them, regardless of the number of mappings performed; there is no dissimilarity to be detected!

Several curves of Pr(γ,p) versus p for different values of γ are shown in Fig. 3. From the figure,

- All curves of Pr(γ,p) start from a value of zero at $p = 1$ (no possibility of detecting dissimilarity on the 1st mapping) and approach unity for large p; as more mappings are performed, dissimilarity is surely to be detected (provided that $\gamma < 1$).

- As γ increases (images become more similar), the curves take longer to reach unity and hence more pixel mappings are required to detect dissimilarity.

- Also noticeable is that all curves quickly reach high probability values, indicating quick dissimilarity detection by using only a few mappings, e.g. Pr(0.4,7) > 0.9. Even when images have good similarity, a few more mappings are required; e.g. Pr(0.8,22) > 0.9.

Fig. 4 shows curves of Pr(γ,p) versus γ at different iso-p (constant p) curves. For any iso-p curve, *DC* decreases with increasing γ. As the value of p increases, the *DC* value also increases at any constant value of γ. This is an informative plot; e.g. it shows that detecting dissimilarity on the 2nd mapping is possible for all $\gamma < 1$; in particular the possibility is 50% for distinct-dissimilar images and decreases as γ increases. It is surprising to observe that even for near-duplicate image pairs ($\gamma \geq 0.99$), such as those shown for the *Leena* images of **Fig. 5**, that the possibility of "getting lucky" and detecting dissimilarity on the 2nd mapping –even though minute (Pr(0.99,2) = 0.1)– nevertheless still exists,

regardless of image size! Note that according to *PMMBI*, 200 mappings are required –on average– to detect dissimilarity for images with this level of similarity.

   Fig. 6 shows curves of *p* versus γ for several *DC* values. For any iso-*DC* curve, *p* increases with increasing γ; the rate becomes larger at higher γ values. Higher *DC* values require more mappings at a given value of γ than lower *DC* values.
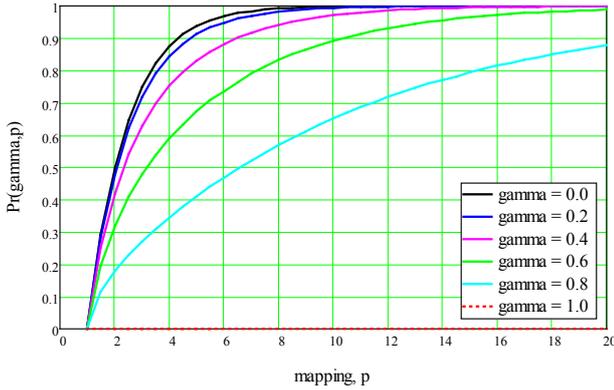

Fig. 6. Highly similar near-duplicate binary images of *Leena* (γ = 0.99). The difference image is also shown.

Fig. 3. Pr(γ, *p*) versus *p* for several iso-γ curves.
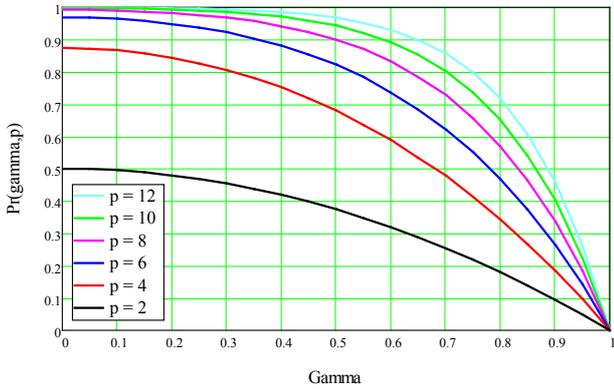
Fig. 4. Pr(γ, *p*) versus γ for several iso-*p* curves.
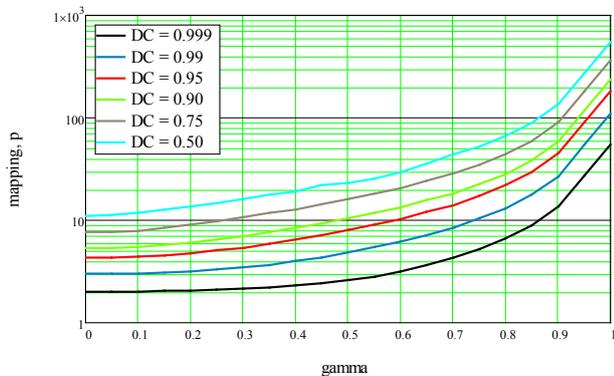
Fig. 5. *p* versus γ for several iso-*DC* curves.

4.4 The Probability Mass Function P_D

   Pr() is a cumulative distribution function in *p* of the probability mass function, $P_D(p,\gamma)$, of the number of mappings required to detect dissimilarity. The by definition [55], $P_D(p,\gamma)$ is obtained from:

$$P_D(p,\gamma) = \Pr(p,\gamma) - \Pr(p-1,\gamma) \qquad (22)$$

Substituting (18) in this equation,

$$P_D(p,\gamma) = 1 - \left(\frac{1}{2}\right)^p \left((1+\gamma)^p + (1-\gamma)^p\right)$$
$$- \left(1 - \left(\frac{1}{2}\right)^{p-1}\left((1+\gamma)^{p-1} + (1-\gamma)^{p-1}\right)\right) \qquad (23)$$

Then simplifying, this becomes,

$$P_D(p,\gamma) = \left(\frac{1}{2}\right)^{p-1}\left((1+\gamma)^{p-1} + (1-\gamma)^{p-1}\right)$$
$$- \left(\frac{1}{2}\right)^p\left((1+\gamma)^p + (1-\gamma)^p\right) \qquad (24)$$

for *p* = 2, 3, … ∞ and 0 ≤ γ < 1. Expanding the terms and rearranging similar terms,

$$P_D(p,\gamma) = \left(\frac{1}{2}\right)^p \left(\frac{(1+\gamma)^p + 2\gamma(1-\gamma)^p + \gamma^2(1+\gamma)^p + (1-\gamma)^p + \gamma^2(1-\gamma)^p - 2\gamma(1+\gamma)^p}{(1-\gamma)(1+\gamma)}\right)$$
$$= \left(\frac{1}{2}\right)^p \frac{\left((1+\gamma)^p + \gamma^2(1+\gamma)^p - 2\gamma(1+\gamma)^p + 2\gamma(1-\gamma)^p + (1-\gamma)^p + \gamma^2(1-\gamma)^p\right)}{(1-\gamma)(1+\gamma)} \qquad (25)$$

Collecting terms and completing the squares,

$$P_D(p,\gamma) = \left(\frac{1}{2}\right)^p \frac{\left((1+\gamma)^p(1+\gamma^2-2\gamma) + (1-\gamma)^p(2\gamma+1+\gamma^2)\right)}{(1-\gamma)(1+\gamma)}$$
$$= \left(\frac{1}{2}\right)^p \frac{\left((1+\gamma)^p(1-\gamma)^2 + (1-\gamma)^p(1+\gamma)^2\right)}{(1-\gamma)(1+\gamma)} \qquad (26)$$

Rearranging and cancelling equivalent terms results in,

$$P_D(p,\gamma) = \frac{(1-\gamma)}{(1+\gamma)}\left(\frac{1}{2}(1+\gamma)\right)^p + \frac{(1+\gamma)}{(1-\gamma)}\left(\frac{1}{2}(1-\gamma)\right)^p$$
$$p = 2, 3, … \infty \text{ and } 0 \le \gamma < 1 \qquad (27)$$

Finally this can be written as,

$$P_D(p,\gamma) = \left(\frac{1}{2}\right)^p \left((1-\gamma)(1+\gamma)^{p-1} + (1-\gamma)^{p-1}(1+\gamma)\right)$$

$$p = 2, 3, \dots \infty \text{ and } 0 \leq \gamma < 1 \quad (28)$$

P_D is a bivariate probability density function in the variables p and γ; once again p is a discrete variable with values $p = 2, 3, \dots, \infty$, and γ is a continuous variable in the range [0,1]. Fig. 7 shows plots of $P_D(p,\gamma)$ as a function of p for values of $\gamma = 0.0$, 0.25, 0.5, 0.75 and 0.9.

For distinct-dissimilar images (D), $\gamma = 0$ and thus,

$$P_D(D,p) = P_D(\gamma=0,p) = \left(\frac{1}{2}\right)^{p-1} \quad (29)$$

It can be seen that most of the weight of $P_D(D, p)$ is at the low values of p concentrated in the first few terms. For example, the first four terms encompasses more than 93% of the total probability mass function of $P_D(D, p)$:

$$\sum_{p=2}^{5} P_D(D,p) = \Pr(D,5) = 0.938 \quad (30)$$

As γ increases, the weight of the $P_D(D, p)$ terms becomes more evenly distributed, as shown in the plots as γ progresses from 0.0 to 0.9.

4.5 P_D is a Probability Density Function

P_D as given by (28) is a probability density function satisfying,

$$\int_0^1 \sum_{p=2}^{\infty} P_D(p,\gamma)d\gamma = 1 \quad (31)$$

which we prove as follows; substituting (28) into (31) yields,

$$\int_0^1 \sum_{p=2}^{\infty} P_D(p,\gamma)d\gamma = \int_0^1 \sum_{p=2}^{\infty} \left(\frac{(1-\gamma)}{(1+\gamma)}\left(\frac{1}{2}(1+\gamma)\right)^p + \frac{(1+\gamma)}{(1-\gamma)}\left(\frac{1}{2}(1-\gamma)\right)^p\right)d\gamma$$

$$= \int_0^1 \left(\frac{(1-\gamma)}{(1+\gamma)}\sum_{p=2}^{\infty}\left(\frac{1}{2}(1+\gamma)\right)^p + \frac{(1+\gamma)}{(1-\gamma)}\sum_{p=2}^{\infty}\left(\frac{1}{2}(1-\gamma)\right)^p\right)d\gamma$$

$$(32)$$

Let $u_1 = \frac{1}{2}(1 + \gamma)$ and $u_2 = \frac{1}{2}(1 - \gamma)$, then the summation terms are of the form Σu^p. Since by definition $0 \leq \gamma < 1$, and hence $|u_1| < 1$ and $|u_2| < 1$ for all γ, then (A2) can be employed. As a result, the first summation term in (32) reduces to,

$$\sum_{p=2}^{\infty}\left(\frac{1}{2}(1+\gamma)\right)^p = \frac{\left(\frac{1}{2}(1+\gamma)\right)^2}{\left(1-\left(\frac{1}{2}(1+\gamma)\right)\right)} \quad (33)$$

Fig. 7. $P_D(\gamma,p)$ as a function of p for different values of γ; from top to bottom $\gamma = 0.0$, 0.25, 0.5, 0.75 and 0.9.

Similarly, using (A2) for the second summation term in (32) produces,

$$\sum_{p=2}^{\infty}\left(\frac{1}{2}(1-\gamma)\right)^p = \frac{\left(\frac{1}{2}(1-\gamma)\right)^2}{\left(1-\left(\frac{1}{2}(1-\gamma)\right)\right)} \quad (34)$$
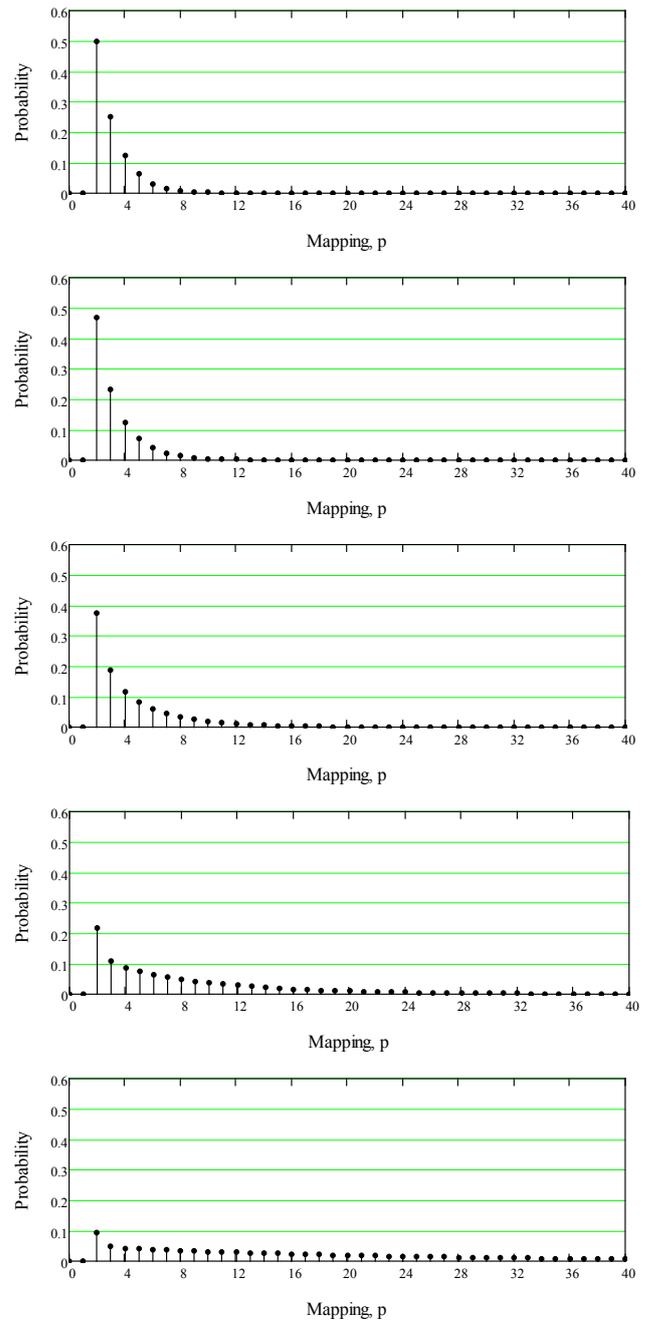
Substituting these two results back in (32) and simplifying,

$$\int_0^1 \sum_{p=2}^{\infty} P_D(p,\gamma) d\gamma = \int_0^1 \left| \frac{(1-\gamma)}{(1+\gamma)} \left(\frac{\left(\frac{1}{2}(1+\gamma)\right)^2}{1-\left(\frac{1}{2}(1+\gamma)\right)} \right) + \frac{(1+\gamma)}{(1-\gamma)} \left(\frac{\left(\frac{1}{2}(1-\gamma)\right)^2}{1-\left(\frac{1}{2}(1-\gamma)\right)} \right) \right| d\gamma$$

$$= \int_0^1 \frac{1}{4} \left| \frac{(1-\gamma)(1+\gamma)}{\frac{1}{2}(1-\gamma)} + \frac{(1+\gamma)(1-\gamma)}{\frac{1}{2}(1+\gamma)} \right| d\gamma \tag{35}$$

Manipulating the algebra we finally obtain the desired result,

$$\int_0^1 \sum_{p=2}^{\infty} P_D(p,\gamma) d\gamma = \int_0^1 \frac{1}{2} \left((1+\gamma) + (1-\gamma) \right) d\gamma$$

$$= \int_0^1 1 \, d\gamma \tag{36}$$

$$= 1$$

□

4.6 The Expected Value and Variance of p

In this section we derive the formula for the expected value of p, which has significant meaning as it represents the average number of mappings required to detect dissimilarity. The formula for the variance of p is also derived.

1. The expected value of P_D

By definition of the expected value [55], the expected value of this distribution is then given by,

$$E[p(\gamma)] = \sum_p (p \cdot P_D(p,\gamma)) \tag{37}$$

Hence,

$$E[p(\gamma)] = \sum_{p=2}^{\infty} \left(p \left(\frac{(1-\gamma)}{(1+\gamma)} \left(\frac{1}{2}(1+\gamma) \right)^p + \frac{(1+\gamma)}{(1-\gamma)} \left(\frac{1}{2}(1-\gamma) \right)^p \right) \right) \tag{38}$$

Simplifying,

$$E[p(\gamma)] = \frac{(1-\gamma)}{(1+\gamma)} \sum_{p=2}^{\infty} \left(p \left(\frac{1}{2}(1+\gamma) \right)^p \right)$$

$$+ \frac{(1+\gamma)}{(1-\gamma)} \sum_{p=2}^{\infty} \left(p \left(\frac{1}{2}(1-\gamma) \right)^p \right) \tag{39}$$

The summation terms have the form $\Sigma p x^p$. Using (A7), the first summation term in (39) reduces to,

$$\sum_{p=2}^{\infty} p \left(\frac{1}{2}(1+\gamma) \right)^p = \frac{\frac{1}{2}(1+\gamma)}{\left(\frac{1}{2}(1+\gamma) - 1 \right)^2} - \frac{1}{2}(1+\gamma)$$

$$= \frac{1}{2}(1+\gamma) \left(\frac{1}{\left(\frac{1}{2}(\gamma-1) \right)^2} - 1 \right) \tag{40}$$

Similarly, using (A7) for the second summation term in (39) produces,

$$\sum_{p=2}^{\infty} p \left(\frac{1}{2}(1-\gamma) \right)^p = \frac{\frac{1}{2}(1-\gamma)}{\left(\frac{1}{2}(1-\gamma) - 1 \right)^2} - \frac{1}{2}(1-\gamma)$$

$$= \frac{1}{2}(1-\gamma) \left(\frac{1}{\left(-\frac{1}{2}(\gamma+1) \right)^2} - 1 \right) \tag{41}$$

Substituting these two results back in (39) and simplifying produces,

$$E[p(\gamma)] = \frac{(1-\gamma)}{(1+\gamma)} \left(\frac{1}{2}(1+\gamma) \right) \left(\frac{1}{\left(\frac{1}{2}(\gamma-1) \right)^2} - 1 \right)$$

$$+ \frac{(1+\gamma)}{(1-\gamma)} \left(\frac{1}{2}(1-\gamma) \right) \left(\frac{1}{\left(-\frac{1}{2}(\gamma+1) \right)^2} - 1 \right) \tag{42}$$

This can be simplified to,

$$E[p(\gamma)] = \left(\frac{1}{2}\right)(1-\gamma)\left(\frac{1}{\left(\frac{1}{2}(\gamma-1)\right)^2} - 1\right)$$
$$+ \left(\frac{1}{2}\right)(1+\gamma)\left(\frac{1}{\left(-\frac{1}{2}(\gamma+1)\right)^2} - 1\right) \quad (43)$$

Further simplification,

$$E[p(\gamma)] = \frac{-1}{\frac{1}{2}(\gamma-1)}\left(1 - \left(\frac{1}{2}(\gamma-1)\right)^2\right)$$
$$+ \frac{1}{\frac{1}{2}(\gamma+1)}\left(1 - \left(-\frac{1}{2}(\gamma+1)\right)^2\right) \quad (44)$$

$$E[p(\gamma)] = \frac{-2}{(\gamma-1)}\left(1 - \left(\frac{1}{2}\right)^2(\gamma-1)^2\right)$$
$$+ \frac{2}{(\gamma+1)}\left(1 - \left(-\frac{1}{2}\right)^2(\gamma+1)^2\right) \quad (45)$$

Multiplying through and simplifying,

$$E[p(\gamma)] = \frac{-2}{(\gamma-1)} + \frac{1}{2}(\gamma-1) + \frac{2}{(\gamma+1)} - \frac{1}{2}(\gamma+1)$$
$$= \frac{-2}{(\gamma-1)} + \frac{2}{(\gamma+1)} - 1 \quad (46)$$

or,

$$E[p(\gamma)] = \frac{2}{(\gamma-1)(\gamma+1)}(-(\gamma+1) + (\gamma-1)) - 1$$
$$= \frac{2(-2)}{(\gamma-1)(\gamma+1)} - 1 \quad (47)$$

which finally simplifies to,

$$E[p(\gamma)] = \frac{4}{1-\gamma^2} - 1 \qquad 0 \le \gamma \le 1 \quad (48)$$

This equation gives the expected value of p, $E[p(\gamma)]$, which is the average number of mappings required to detect dissimilarity at any given γ, which will be denoted by p^*,

$$p^* = \frac{4}{1-\gamma^2} - 1 \qquad 0 \le \gamma \le 1 \quad (49)$$

When the images are highly similar and near duplicate, γ is close to unity, and thus $(\gamma + 1) \cong 2$. Then from (49), p^* can be approximated by,

$$p^* = E(\gamma) \approx \frac{2}{1-\gamma} \qquad \gamma \cong 1 \quad (50)$$

Using this approximation produces an error < 2.5% for $\gamma > 0.705$, and an error < 1% for $\gamma > 0.809$.

When the images are highly dissimilar and γ is small, then from (49), p^* can be approximated by,

$$p^* = E(\gamma) \approx 3 \qquad \gamma \cong 0 \quad (51)$$

Using this approximation produces an error < 2.5% for $\gamma < 0.137$ and an error < 1% for $\gamma < 0.087$. Eq. (51) also implies that this is the lowest possible expected mapping. This should not be incorrectly misinterpreted that 3 point mapping is the lowest possible number of mappings required to detect dissimilarity; it was already shown above that 2 mappings are possible to detect dissimilarity.

2. The Variance of p

By definition, the variance of p is [55],

$$V[p] = E[p^2] - (E[p])^2 \quad (52)$$

For the first term, $E[p^2]$, we proceed as follows,

$$E[p^2(\gamma)] = \sum_p (p^2 \cdot P_D(p,\gamma))$$
$$= \sum_{p=2}^{\infty}\left(p^2\left(\frac{(1-\gamma)}{(1+\gamma)}\left(\frac{1}{2}(1+\gamma)\right)^p + \frac{(1+\gamma)}{(1-\gamma)}\left(\frac{1}{2}(1-\gamma)\right)^p\right)\right)$$
$$= \frac{(1-\gamma)}{(1+\gamma)}\sum_{p=2}^{\infty}\left(p^2\left(\frac{1}{2}(1+\gamma)\right)^p\right) + \frac{(1+\gamma)}{(1-\gamma)}\sum_{p=2}^{\infty}\left(p^2\left(\frac{1}{2}(1-\gamma)\right)^p\right) \quad (53)$$

The summation terms have the form $\Sigma p^2 x^p$. Using (A11), the first summation term in (53) reduces to,

$$\sum_{p=2}^{\infty} p^2\left(\frac{1}{2}(1+\gamma)\right)^p = \frac{1}{2}(1+\gamma)\left(\frac{1+\frac{1}{2}(1+\gamma)}{\left(1-\frac{1}{2}(1+\gamma)\right)^3} - 1\right) \quad (54)$$

Similarly, using (A11) for the second summation term in (53) produces,

$$\sum_{p=2}^{\infty} p^2\left(\frac{1}{2}(1-\gamma)\right)^p = \frac{1}{2}(1-\gamma)\left(\frac{1+\frac{1}{2}(1-\gamma)}{\left(1-\frac{1}{2}(1-\gamma)\right)^3} - 1\right) \quad (55)$$

Thus, substituting back in (53) results in,

$$E[p^2(\gamma)] = \frac{(1-\gamma)}{(1+\gamma)}\left(\frac{1}{2}(1+\gamma)\left(\frac{1+\frac{1}{2}(1+\gamma)}{\left(1-\frac{1}{2}(1+\gamma)\right)^3}-1\right)\right)$$
$$+\frac{(1+\gamma)}{(1-\gamma)}\left(\frac{1}{2}(1-\gamma)\left(\frac{1+\frac{1}{2}(1-\gamma)}{\left(1-\frac{1}{2}(1-\gamma)\right)^3}-1\right)\right) \quad (56)$$

Multiplying the terms out and rearranging produces,

$$E[p^2(\gamma)] = \frac{1}{2}\left(\,(1-\gamma)\left(\frac{1+\frac{1}{2}(1+\gamma)}{\left(1-\frac{1}{2}(1+\gamma)\right)^3}-1\right)\right.$$
$$\left.+\left((1+\gamma)\left(\frac{1+\frac{1}{2}(1-\gamma)}{\left(1-\frac{1}{2}(1-\gamma)\right)^3}-1\right)\right)\right) \quad (57)$$

or,

$$E[p^2(\gamma)] = \frac{1}{2}\left(\,(1-\gamma)\left(\frac{1+\frac{1}{2}(1+\gamma)}{\left(\frac{1}{2}(1-\gamma)\right)^3}-1\right)\right.$$
$$\left.+\left((1+\gamma)\left(\frac{1+\frac{1}{2}(1-\gamma)}{\left(\frac{1}{2}(1+\gamma)\right)^3}-1\right)\right)\right) \quad (58)$$

Further simplification results in,

$$E[p^2(\gamma)] = 4\left(\frac{1}{(1-\gamma)^2}\left(1+\frac{1}{2}(1+\gamma)-\left(\frac{1}{2}(1-\gamma)\right)^3\right)\right)$$
$$+4\left(\frac{1}{(1+\gamma)^2}\left(1+\frac{1}{2}(1-\gamma)-\left(\frac{1}{2}(1+\gamma)\right)^3\right)\right) \quad (59)$$

This equation finally simplifies to,

$$E[p^2(\gamma)] = \frac{11+22\cdot\gamma^2-\gamma^4}{(1-\gamma^2)^2} \quad (60)$$

Hence, substituting this result and $E[p]$ back into (52) produces,

$$V[p] = \frac{11+22\cdot\gamma^2-\gamma^4}{(1-\gamma^2)^2}-\left(\frac{4}{1-\gamma^2}-1\right)^2 \quad (61)$$

Simplifying we arrive at the final result,

$$V[p] = \frac{2(8\cdot\gamma^2-\gamma^4+1)}{(1-\gamma^2)^2} \quad (62)$$

When the images are highly dissimilar and γ is small, then,

$$V[p(\gamma)] \approx 2 \quad (63)$$

Using this approximation produces an error < 2.5% for γ < 0.050 and an error < 1% for γ < 0.031.

A plot of $E(\gamma)$ and $V(\gamma)$ is shown in Fig. 8. $E(\gamma)$ increases slowly for low γ, but increases rapidly at higher γ values. $V(\gamma)$ also increases as γ increases, but at a quicker rate than $E(\gamma)$.

4.7 Measuring similarity with *PMMBI*

By measuring how quickly dissimilarity can be detected between images, i.e. how many mappings are required for detection and using *PMMBI*, the similarity (γ) between images can be estimated to a good degree. In such cases, several dissimilarity detection trials should be repeated and the mean value can be used as the value of p^*. Then (49) can be used to estimate the amount of similarity, γ, between the images (see [56]). As a result, similarity can be measured and estimated to a good degree without the need to scan the entire images. This also implies that matching can be performed quickly regardless of image size.
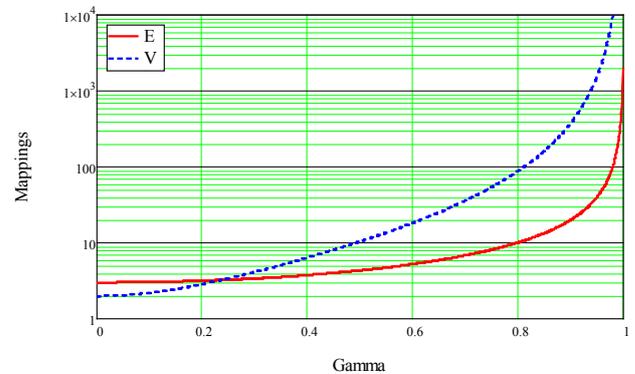

Fig. 8. Plots of $E(\gamma)$ and $V(\gamma)$.

5 Discussion

Fig. 9 shows the images of the *Reduced KU-ME128B* binary image set that was used for testing the probability model. The set consists of 12 128x128 binary images of different scenes. The similarity values between all image pairs for this set are in the range of $0.002 \leq \gamma \leq 0.669$, with a mean value of 0.231 and a standard deviation of 0.161. Every image was matched to every other image and the number of mappings, *MDN*, required for detection was recorded. Since with random mapping every mapping trial produces a different dissimilarity detection outcome as measured by *MDN*, matching of each pair of images was repeated 1,000 times to obtain consistent and more accurate statistical results. The resulting $MDN_{0.50}$, $MDN_{0.90}$, $MDN_{0.99}$ and $MDN_{0.999}$ are plotted as a function of γ in Fig. 10. The theoretical iso-*DC* curves for 0.50, 0.90, 0.99 and 0.99 are also plotted in the figure for comparison. **Table 1** summarizes the dissimilarity detection mapping error statistics. Examining the discrepancy between *PMMBI*, the model's prediction, with the empirical results obtained, we observe the following:

- All MDN_{DC} values have a high correlation (0.839 – 0.989) with the theoretical *DC* values as predicted by *PMMBI* as given in (18).

- The mean mapping error between predicted and empirical data is very small with less than one mapping for all *DC* values, except for $MDN_{0.999}$ which has a higher mean error value of ~1.5 mappings, which is expected due to the larger expected value for *p* at higher *DC* values.

The results exemplify the high accuracy of the model in predicting how fast dissimilarity can be detected.

Fig. 9. The *Reduced KU-ME128* binary image set.
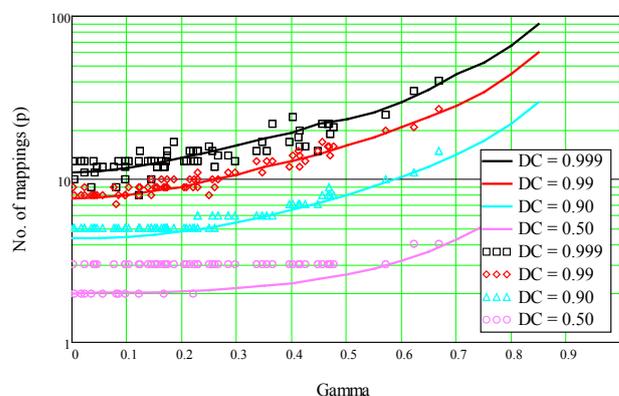

Fig. 10. Plots of MDN_{DC} versus γ for *DC* = 0.50, 0.90, 0.99 and 0.999. The theoretical *DC* curves are also shown.

6 Conclusion

In this paper we have presented a probabilistic model, called the *Probabilistic Matching Model for Binary Images* (*PMMBI*), for the quick detection of dissimilarity between binary images. The model is based on randomly mapping pixels between images. The model predicts the probability of detecting dissimilarity between binary images as a function of the similarity between images and the number of mappings between them. The model shows that dissimilarity can be detected fairly quickly when the images are highly dissimilar, requiring only a few mappings between the images. As the images become more similar, more mappings are required to detect similarity, but still only a small fraction compared to processing the entire images. Even near-duplicate images, where more than 99.5% of the image content is similar, the model shows that on average only 200 mappings are required to detect dissimilarity, regardless of image size! The model's

Table 1: *Mapping Results Statistics for the Reduced KU-ME128 binary image set*				
		Mapping Error		
DC	**Correlation**	**Max**	**Mean**	**Std. dev.**
0.500	0.839	1.000	0.629	0.391
0.900	0.989	2.185	0.455	0.353
0.990	0.986	2.464	0.675	0.530
0.999	0.956	6.868	1.502	1.359

invariance to image size is a unique feature of the model that gives it its strength, particularly when the images are huge. Testing with real images produced dissimilarity detection results in agreement with that predicted by the model, showing the accuracy of the model.

Our future work will focus on showing how *PMMBI* can be used efficiently for template matching and image registration, even in the presence of noise. We will also focus on developing a similar probabilistic model for greyscale images.

References:

[1] Gong, M., Zhao, S., Jiao, L., Tian, D., and Wang. S., "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information." *IEEE Transactions on Geoscience and Remote Sensing*, 52, 7, 2014, pp. 4328-4338.

[2] Oliveira, F., and Tavares, J., "Medical image registration: a review." *Computer methods in biomechanics and biomedical engineering*, 17, 2, 2014, pp. 73-93.

[3] Ahuja, K. and Tuli, P. "Object recognition by template matching using correlations and phase angle method". *International Journal of Advanced Research in Computer and Communication Engineering*, 2, 3, 2013, pp. 1368-1373.

[4] Lewis, J., "Fast Template Matching", *Vision Interface*, 1995, pp. 120-123.

[5] Gong, Y., Lazebnik, S., Gordo, A., and Perronnin, F., "Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35, 12, 2013, pp. 2916-2929.

[6] Park, U., Park, J. and Jain, A. "Robust keypoint detection using higher-order scale space derivatives: application to image retrieval". *IEEE Signal Processing Letters*, 21, 8, 2014, pp.962-965.

[7] Korytkowski, M., Rutkowski, L., and Scherer, R., "Fast image classification by boosting fuzzy classifiers." *Information Sciences*, 327, 2016, pp. 175-182.

[8] Canty, M., *Image analysis, classification and change detection in remote sensing: with algorithms for ENVI/IDL and Python*. Crc Press, 2014.

[9] Hii, A., Hann, C., Chase, J. and Van Houten, E., "Fast normalized cross correlation for motion tracking using basis functions",

Computer Methods and Programs in Biomedicine, 8, 2, 2006, pp. 144–156.

[10] Luo, J. and Konofagou, E., "A Fast Normalized Cross-Correlation Calculation Method for Motion Estimation", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57, 6, 2010, pp. 1347-1357.

[11] Lin, C. and Chen, C., "Fast normalized cross correlation for defect detection", *Pattern Recognition Letters*, 24, 15, 2003, pp. 2625–2631.

[12] Lowe, D., Object recognition from local scale-invariant features. In *Computer vision*, the proceedings of the *seventh IEEE international conference on*, Vol. 2, pp. 1150-1157.

[13] Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. Speeded-up robust features (SURF). *Computer vision and image understanding*, 110, 3, 2008, pp. 346-359.

[14] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G., "ORB: An efficient alternative to SIFT or SURF". In *Computer Vision* (*ICCV*), *2011 IEEE international conference* on, 2011, pp. 2564-2571.

[15] Anuta, P., "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques", *IEEE Trans. on Geoscience Electronics*, GE-8, N 4, 1970, pp. 353-368.

[16] Barnea, D. and Silverman, H., "A Class of Algorithms for Fast Digital Image Registration". *IEEE Trans. on Computers*, Vol. c-21, N 2, 1972, pp.179-186.

[17] Li, R., Zeng, B. and Liou, M., "A New Three-Step Search Algorithm for Block Motion Estimation", *IEEE Transactions on Circuits and Systems for Video Technology*, 4, 4, 1994, pp. 438-442.

[18] Zhang, K., Lu, J., Lafruit, G., Lauwereins, R., and Van Gool, L. "Robust stereo matching with fast normalized cross-correlation over shape-adaptive regions", *16th IEEE International Conference on Image Processing* (*ICIP*), 2009, pp. 2357-2360.

[19] Kim, H., "Rotation-discriminating template matching based on Fourier coefficients of radial projections with robustness to scaling and partial occlusion", *Pattern Recognition*, 43, 3, 2010, pp. 859-872.

[20] Mukherji, S., "Fast Algorithms for Binary Cross-correlation", *proceedings of Geoscience and Remote Sensing Symposium*, 1, 2005, pp. 4-11.

[21] Tang, F. and Tao, H., "Fast multi-scale template matching using binary features", 8th

IEEE Workshop on Applications of Computer Vision, 2007, pp. 36-39.

[22] Liu, M. and Li, L., "Cross-correlation based binary image registration for 3D palmprint recognition", *11th IEEE International Conference on Signal Processing*, 2012, pp. 1597-1600.

[23] Mattoccia, S., Tombari, F, Di-Stefano, L., "Reliable rejection of mismatching candidates for efficient ZNCC template matching", *15th IEEE International Conference on Image Processing*, 2008, 849-852.

[24] Yoo, J., Choi, B., and Choi, H., "1-D fast normalized cross-correlation using additions". *Digital Signal Processing*, 20, 5, 2010, pp. 1482-1493.

[25] Choi, M. and Kim, W., "A novel two stage template matching method for rotation and illumination invariance", *Pattern Recognition*, 35, 1, 2002, pp. 119-129.

[26] Lin, Y, and Chen, C., "Template matching using the parametric template vector with translation, rotation and scale invariance", *Pattern Recognition*, 41, 7, 2008, pp. 2413-2421.

[27] Chen, Y., Hung, Y., and Fuh, C., "Fast Block Matching Algorithm Based on the Winner-Update Strategy", *IEEE Transactions on Image Processing*, 10, 8, 2001, pp. 1212-1222.

[28] Fouda, Y. "A Robust Template Matching Algorithm Based on R educing Dimensions", *Journal of Signal and Information Processing*, 6, 2015, pp. 109-122.

[29] Vassiliadis, S., Hakkennes, E., Wong, J., and Pechanek, G., "The sum-absolute-difference motion estimation accelerator", *Proceedings of the IEEE Euromicro Conference*, 2, 199 8, pp. 559-566.

[30] Guevorkian, D., Launiainen, A., Liuha, P., and Lappalainen, V., "Architectures for the sum of absolute differences operation", *IEEE Workshop on Signal Processing Systems*, 2002, pp. 57-62.

[31] Vanne, J., Aho, E., Hämäläinen, T. and Kuusilinna, K., "A High-Performance Sum of Absolute Difference Implementation for Motion Estimation", *IEEE Transactions on Circuits and Systems for Video Technology*, 16, 7, 2006, pp. 876-883.

[32] Niitsuma, H., and Maruyama, T., "Sum of Absolute Difference Implementations for image processing on FPGAs", *International Conference on Field Programmable Logic and Applications*, 2010, pp. 167-170.

[33] Wong, S., Vassiliadis, S., and Cotofana, S., "A Sum of Absolute Differences Implementation in FPGA Hardware", *Proceedings of the 28th Euromicro Conference*, 2002, pp. 183-188.

[34] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P., "Multimodality Image Registration by Maximization of Mutual Information", *IEEE Transactions on Medical Imaging*, 16, 2, 1997, pp. 187-198.

[35] Tomaževič, D., Likar, B. and Pernuš, F., "Multi-Feature Mutual Information Image Registration", *Image Anal Stereol*, 31, 2012 , pp. 43-53.

[36] Pluim, J., Maintz, A., and Viergever, M., "Mutual-Information-Based Registration of Medical Images: A Survey", *IEEE Transactions on Medical Imaging*, 22, 8, 2003.

[37] Mustafa, A. and Ganter, M., "An Efficient Image Registration Method by Minimizing Intensity Combinations", *Research in Computer and Robot Vision*, Archibald, C. and Kwok, P. (Eds.), World Scientific Press, Singapore, 1995, pp. 247-268.

[38] Baudrier, E., Nicolier, F., Millon, G. and Ruan, S., "Binary-image comparison method with local-dissimilarity quantification", *Pattern Recognition*, 41, 2008, pp. 1461-1478.

[39] Vidal, J. and Crespo, J., "Sets Matching in Binary Images Using Mathematical Morphology", *International Conference of the Chilean Computer Science Society*, 2008, pp. 110-115.

[40] Teshome, M., Zerubabe, L. and Yoon, K., "A Simple Binary Image Similarity Matching Method Based on E xact Pixel Matching", *International Conference on Computer Engineering and Applications*, 2009, pp. 12-15.

[41] Mustafa, A., "Probabilistic Model for Quick Detection of Dissimilar Binary Images". *Journal of Electronic Imaging*, 24, 5, 2015, pp. 24-53.

[42] Mustafa, A., "A Modified Hamming Distance Measure for Quick Rejection of Dissimilar Binary Images". *International Conference on Computer Vision and Image Analysis*, 2015.

[43] Mustafa, A., "A Probabilistic Binary Similarity Distance for Quick Image Matching". *IET Journal on Image Processing*, to be submitted.

[44] Zhang, L., Zeng, Z., and Ji, Q., "Probabilistic Image Modeling With an Extended Chain Graph for Human Activity Recognition and Image Segmentation". *IEEE Transactions on Image Processing*, 20, 9, 2011, pp. 2401- 2413.

[45] Yi, W., Chen, Y., Tang, H. and Deng, L., "Experimental research on ur ban road

extraction from high-resolution RS images using Probabilistic Topic Models". *IEEE International Geoscience and Remote Sensing Symposium*, 2010, pp. 445-448.

[46] Yiu, W. (). A fast probabilistic method for vehicle detection and tracking with an explicit contour model. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR. 2005.

[47] Ayromlou, M., Vincze, M. and Ponweiser, W., "Probabilistic matching of image- to model-features for real-time object tracking". *Proceedings 16th International Conference on Pattern Recognition*, 2002, pp. 692- 695.

[48] Risholm, P., Fedorov, A., Pursley, J., Tuncali, K., et al. "Probabilistic non-rigid registration of prostate images: Modeling and quantifying uncertainty". *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2011, pp. 553-556

[49] Wen, C., Guo, C., and Wen, C., "Multiresolution Image Fusion Algorithm Based on Probabilistic Model". *The Sixth World Congress on Intelligent Control and Automation*, 2006, pp. 10398-10402.

[50] Jin, R., and Hauptmann, A., "Using a probabilistic source model for comparing images", *International Conference on Image Processing*, 2002, pp. 941- 944.

[51] Yamaguchi, T., and Maruyama, M., "Image categorization by a cl assifier based on probabilistic topic model". *19th International Conference on Pattern Recognition*, 2008, pp.1-4.

[52] Zhang, R., Zhang, Z., Li, M., Ma, W. et al., "A probabilistic semantic model for image annotation and multimodal image retrieval". *Tenth IEEE International Conference on Computer Vision*, 2005, pp. 846-851.

[53] Ning, Z., Cheung, W., Guoping, Q. and Xiangyang, X., "A Hybrid Probabilistic Model for Unified Collaborative and Content-Based Image Tagging". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 7, 2011, pp. 1281-1294.

[54] Mustafa, A. "A Framework for Quick Rejection of Dissimilar Binary Images". *International Journal of Signal Processing Systems*, 1, 2, 2013, pp. 237-243.

[55] Scheaffer, R., a nd McClave, J., *Probability and statistics for Engineers*, Dixbury press, Fourth Ed., Calif., U.S.A., 1995.

[56] Mustafa, A., "Quick Probabilistic Binary Image Matching: Changing the Rules of the Game". *Proc. SPIE 9971, Applications of Digital Image Processing XXXIX*, 997112, 2016.

6 Appendix

In this section, we provide the proofs for several infinite series results used in the theoretical derivations of our work. The following infinite series is a geometric series with constant 1 and ratio x with sum given by,

$$\sum_{p=0}^{\infty} x^p = 1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x} \ , \ |x| < 1$$

(A1)

As long as $|x| < 1$, then the sum converges.

A.1 The sum of an infinite series of the form x^p starting at $p = 2$

The sum of an infinite series of the form x^p starting at $p = 2$ is given by,

$$\sum_{p=2}^{\infty} x^p = \frac{x^2}{1-x} \qquad |x| < 1 \quad \text{(A2)}$$

Proof: Since,

$$\sum_{p=2}^{\infty} x^p = x^2 + x^3 + x^4 + \cdots \qquad \text{(A3)}$$

Adding and subtracting $(1 + x)$,

$$\sum_{p=2}^{\infty} x^p = (1+x) + x^2 + x^3 + x^4 + \cdots - (1+x)$$

$$= (1 + x + x^2 + x^3 + x^4 + \cdots) - (1+x)$$

(A4)

Recognizing the first term as (A2) produces,

$$\sum_{p=2}^{\infty} x^p = \frac{1}{1-x} - (1+x) \qquad \text{(A5)}$$

Factoring out $(1 - x)$ in the denominator and recognizing the difference in square term,

$$\sum_{p=2}^{\infty} x^p = \frac{1}{1-x}\left(1 - (1+x)(1-x)\right)$$

$$= \frac{1}{1-x}\left(1 - (1-x^2)\right)$$

(A6)

The summation can then be finally stated as,

$$\sum_{p=2}^{\infty} x^p = \frac{x^2}{1-x} \qquad \text{(A2)}$$

□

A.2 The sum of an infinite series of the form px^p starting at $p = 2$

The sum of an infinite series of the form px^p starting at $p = 2$ is,

$$\sum_{p=2}^{\infty} \left(p \cdot x^p\right) = \frac{x}{(x-1)^2} - x \qquad |x| < 1 \quad \text{(A7)}$$

Proof: Since,

$$\sum_{p=2}^{\infty}\left(p \cdot x^p\right) = 2x^2 + 3x^3 + 4x^4 + \cdots$$

$$= x + 2x^2 + 3x^3 + 4x^4 + \cdots - x$$

$$= x(1 + 2x + 3x^2 + 4x^3 + \cdots) - x \qquad \text{(A8)}$$

$$= x\frac{d}{dx}(1 + x + x^2 + x^3 + x^4 + \cdots) - x$$

Writing the term in brackets as a su mmation produces,

$$\sum_{p=2}^{\infty}\left(p \cdot x^p\right) = x\frac{d}{dx}\left(\sum_{p=0}^{\infty} x^p\right) - x \qquad \text{(A9)}$$

Recognizing the summation appearing on the right-hand side of this equation as (A1) produces,

$$\sum_{p=2}^{\infty}\left(p \cdot x^p\right) = x\frac{d}{dx}\left(\frac{1}{1-x}\right) - x$$

$$= x\left(\frac{1}{(x-1)^2}\right) - x \qquad \text{(A10)}$$

This finally produces,

$$\sum_{p=2}^{\infty}\left(p \cdot x^p\right) = \frac{x}{(x-1)^2} - x \qquad |x| < 1 \quad \text{(A7)}$$

□

A.3 The sum of an infinite series of the form $p^2 x^p$ starting at $p = 2$

The sum of an infinite series of the form $p^2 x^p$ starting at $p = 2$ is,

$$\sum_{p=2}^{\infty}\left(p^2 \cdot x^p\right) = x\left(\frac{1+x}{(1-x)^3} - 1\right) \qquad |x| < 1 \quad \text{(A11)}$$

Proof: Since,

$$\sum_{p=2}^{\infty}\left(p^2 \cdot x^p\right) = 4x^2 + 9x^3 + 16x^4 + \cdots \qquad \text{(A12)}$$

$$= (x + 4x^2 + 9x^3 + 16x^4 + \cdots) - x$$

Let the first term on the RHS be S, then,

$$S = x + 4x^2 + 9x^3 + 16x^4 + \cdots \qquad \text{(A13)}$$

$$xS = x^2 + 4x^3 + 9x^4 + 16x^5 + \cdots \qquad \text{(A14)}$$

Subtracting these two equations and factoring out x produces,

$$(1-x)S = x + 3x^2 + 5x^3 + 7x^4 + \cdots$$

$$= x(1 + 3x + 5x^2 + 7x^3 + \cdots) \qquad \text{(A15)}$$

Dividing both sides by x,

$$\frac{(1-x)}{x} S = 1 + 3x + 5x^2 + 7x^3 + \cdots \qquad \text{(A16)}$$

This can be rewritten as,

$$\frac{(1-x)}{x} S = (2 + 4x + 6x^2 + 8x^3 + \cdots)$$

$$- (1 + x + x^2 + x^3 + \cdots) \qquad \text{(A17)}$$

By algebraic manipulation, calculus and (A1), the sum of the first term of (A17) becomes,

$$2 + 4x + 6x^2 + 8x^3 + \cdots = 2(1 + 2x + 3x^2 + 4x^3 + \cdots)$$

$$= 2\frac{d}{dx}(1 + x + x^2 + x^3 + \cdots)$$

$$= \frac{2}{(1-x)^2} \qquad \text{(A18)}$$

The second term of (A17) is given by (A1). Hence,

$$\frac{(1-x)}{x} S = \frac{2}{(1-x)^2} - \frac{1}{1-x} \qquad \text{(A19)}$$

$$S = \frac{x}{1-x}\left(\frac{2}{(1-x)^2} - \frac{1}{1-x}\right) \qquad \text{(A20)}$$

which simplifies to,

$$S = \frac{x(1+x)}{(1-x)^3} \qquad \text{(A21)}$$

Finally substituting this equation back into (A12) produces,

$$\sum_{p=2}^{\infty}\left(p^2 \cdot x^p\right) = \frac{x(1+x)}{(1-x)^3} - x \qquad \text{(A22)}$$

This finally produces,

$$\sum_{p=2}^{\infty}\left(p^2 \cdot x^p\right) = x\left(\frac{1+x}{(1-x)^3} - 1\right) \qquad \text{(A11)}$$

□