An Image Size Invariant Method for Quick Detection of Dissimilar Binary Images

Adnan A. Y. Mustafa, *Member, IAENG*

***Abstract*— In this paper we present an image size invariant method for quick detection of dissimilar binary images. The method is based on a *Probabilistic Matching Model* (*PMM*) for binary image matching. Using the model, the probability of matching dissimilar image pairs can be predicted, as a function of the number of points mapped between two images and the amount of similarity between them. The model tells us that by matching few points between two images, we can determine dissimilar images with high confidence. For example, if images are distinct-dissimilar, i.e., completely different, only 8 points need to be mapped to arrive at a 90% successful detection rate, 11 points need to be mapped for a 99% confidence detection rate and only 15 points need to be mapped for a 99.9% confidence detection rate. If the images are not distinct-dissimilar and the images have some similarity between them, then more points need to be matched; depending on the amount of similarity between the images. The model is image size invariant and hence images of any sizes will produce the same high confidence levels with only a limited number of points. As a result, this method does not suffer from the image size handicap that current methods suffer from. We report on tests conducted on real images of different sizes to show the validity of our model.**

***Index Terms*— binary images, image mapping, image matching and probabilistic model.**

I. Introduction

IMAGE matching rises frequently in the field of image analysis under many topics such as, image registration, template matching, image retrieval, image classification, etc. These methods are either feature-based methods that rely on some method of extracting image features and then matching the extracted features, or area-based methods (also referred to as direct or intensity methods) that are based on comparing image intensity values directly without any feature extraction. When dealing with binary images, and due to the fact that they have only two intensity values resulting in a limited amount of scene detail, feature-based methods become difficult to employ, and area-based methods become the method of choice.

Binary image matching is usually accomplished by calculating the cross-correlation between the images [1] or simply by subtracting the two images [2]. These methods, as well as the majority –if not all– of area-based methods require some type of similarity operation to be applied to the whole image. Hence, these methods are image size dependent which implies that as image size increases, more processing time is required. With 20 Mega-pixel images common and 50 Mega-pixel images becoming easily producible with recently introduced inexpensive digital cameras or even mobile phones, current matching methods can become quite slow processing such large images, even with today's fast computers.

In this paper we present an image size invariant method for quick detection of dissimilar binary images. The method is based on a *Probabilistic Matching Model* (*PMM*) for binary image matching [3]. Using the model, the probability of matching dissimilar image pairs can be predicted, as a function of the number of points mapped between two images and the amount of similarity between them. The model tells us that by matching few points between two images, we can determine dissimilar images with high confidence. For example, if images are distinct-dissimilar, i.e., completely different, only 8 points need to be mapped to arrive at a 90% successful detection rate, 11 points need to be mapped for a 99% confidence detection rate and only 15 points need to be mapped for a 99.9% confidence detection rate. If the images are not distinct-dissimilar and the images have some similarity between them, then more points need to be matched; depending on the amount of similarity between the images. The model is image size invariant and hence images of any sizes will produce the same high confidence levels with only a limited number of points. As a result, this method does not suffer from the image size handicap that current methods suffer from. We report on tests conducted on real images of different sizes to show the validity of our model.

This paper is organized as follows: section II points out related literature, section III reviews the binary image similarity measure used as a reference in our work, section IV presents binary image mappings and how they can simplify binary matching, section V presents the main theme of this paper and presents the theory of the probabilistic matching model, section VI presents results of tests conducted on real images and their agreement with the theoretical model and section VII finally concludes our findings and discusses where our future research is directed.

II. Related Literature

Cross-correlation is the most widely used method for image matching. Many techniques have attempted to improve using cross-correlation; Lebrl [4] counted the change in mismatched pixels as the template sweeps the image. Lewis [5] used pre-computed tables containing the integral of the image and the image^2 search window. Mattoccia et al. [6] presented a zero-mean normalized cross-correlation to reduce the computational cost.

Other area-based methods have also been developed based on a variety of principles. For example, Mustafa et al. [7] matched images by minimizing the image intensity combinations with excellent results. However, the method is

Manuscript received December 8, 2014; revised February 1, 2015.

Adnan A. Mustafa is with the Department of Mechanical Engineering, Kuwait University, P. O. Box 5969-Safat, Kuwait 13060 (phone: (+965) 24987117, Fax:(+965) 24847131; E-mail: adnan.mustafa@ku.edu.kw).

not suited for binary images. Baudrier et al. [8] [9] used local-dissimilarity map (LDMap) produced from the Hausdorff distance. Tang et al. [10] represented the template as a linear combination of a small number of Haar-like binary features. Vidal et al. [11] used mathematical morphology while Sleit et. al. [12] used chain-codes.

III. Image Closeness with the γ Similarity Measure

In this section, we present our definition of similarity between images, and how we categorize images based on it. We then introduce the γ similarity measure that is used in our work as a reference for image similarity.

A. Similar and Dissimilar Images

The closeness between two images is based on a pixel-to-pixel comparison. Image closeness is categorized as either *similar* (*S*) or *dissimilar* (*R*);

- *Similar* images: the two images are considered to be the same and are of two types, *exact* (*E*) or *inverse* (*I*). If the two images have exactly the same intensity values then they are *exact*. If the two images have the complement intensity values then they are *inverse*.
- *Dissimilar* images: the two images are not similar and are of two types: *Distinct-dissimilar* (*D*) and *Quasi-dissimilar* (*Q*). Distinct-dissimilar images are images that are completely different whereas Quasi-dissimilar images have concurrences between them.

B. The γ Binary Similarity Measure

There is an abundance of binary similarity measures and distances that have been developed over the last century and can be found in the literature, e.g. [13]. However, for our work we need a measure that can be used with the categorization of image similarity stated above and the *PMM* theory developed below. The *binary similarity measure* (γ) was specifically developed for this purpose [14]. γ is a modification of the Hamming distance that measures the amount of similarity and concurrence between two images. Formally, given two images **u** and **v**, γ is defined as,

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

where ⊕ is the exclusive-or (XOR) operation and $P_o(Z)$ denotes the probability mass function of Z. γ has the range, $0 \le \gamma \le 1$. Then based on γ we have,

- distinct-dissimilar images when $\gamma = 0$,
- quasi-dissimilar images when $0 < \gamma < 1$, and
- similar images when $\gamma = 1$.

In practice, image pairs with $\gamma < 0.01$ are assumed to be $\gamma \approx 0$, and are considered to be distinct-dissimilar image pairs.

IV. Binary Image Mappings

While the complexity of images increases with image size, fortunately the number of possible image mappings does not. In fact, for binary images there are only 15 different possible image mappings, regardless of image size. These image mappings can simplify the matching process and reduce matching time drastically.

Let $u,v \in \{0,1\}$ be two independent random variables that represent the image intensity values of two images. Hence, for binary images, there are 4 possible pixel mappings between any two images,

$$P_1 = \{A, B, C, D\} \qquad (2)$$

where A:0→0, B:0→1, C:1→0 and D:1→1. Let V_1 refer to the possible *image mapping variations* between two binary images. Since the number of binary pixel mappings is finite, the number of possible image mapping variations between any two binary images is also finite –regardless of image size– and is $NV_1 = 15$. These mapping variations, denoted by μ_i, $i = 1 \ldots 15$, along with their tuple sizes and image closeness interpretation are listed in Table I.

V. A Probabilistic Model for Matching

Probabilistic models have been developed and applied to many areas of computer vision, e.g., image comparison [15] and image retrieval [16]. In this section we present a probabilistic model for image matching. Initially we begin by reviewing the probabilistic model for image mapping [3] which is the basis for the matching model.

A. A Probabilistic Model for Binary Image Mapping

Since *u* and *v* are two independent random variables, all pixel mappings have equal probabilities of occurrence. Let τ_m denote a tuple of size *m* such that $\tau_m \in T_{1,m}$, where $T_{1,m}$ denotes the set consisting of binary image mapping variations with size *m*, then $\Pr(\tau_m, p)$ represents the probability of occurrence of an *m*-tuple on the p^{th} mapped pixel and is given by,

$$\Pr(\tau_m, p) = \frac{m!}{4^p}\binom{4}{m} S_m^p \qquad m = 1 \ldots 4 \qquad (3)$$

where,

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \qquad (4)$$

is the combination function and *S* is the *Stirling numbers of the second kind* function,

TABLE I
THE 15 MAPPING VARIATIONS FOR BINARY IMAGES.

μ_i	pixel mappings	tuple size	Image closeness
μ_1	*A*	1	*similar*
μ_2	*B*	1	*similar*
μ_3	*C*	1	*similar*
μ_4	*D*	1	*similar*
μ_5	*A,B*	2	*dissimilar*
μ_6	*A,C*	2	*dissimilar*
μ_7	*A,D*	2	*similar*
μ_8	*B,C*	2	*similar*
μ_9	*B,D*	2	*dissimilar*
μ_{10}	*C,D*	2	*dissimilar*
μ_{11}	*A,B,C*	3	*dissimilar*
μ_{12}	*A,B,D*	3	*dissimilar*
μ_{13}	*A,C,D*	3	*dissimilar*
μ_{14}	*B,C,D*	3	*dissimilar*
μ_{15}	*A,B,C,D*	4	*Dissimilar*

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

$$S_m^p = \frac{1}{m!}\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}j^p \qquad 0 \le m \le p \qquad (5)$$

B. A Probabilistic Model for Binary Image Matching

Integrating the binary image mapping model with the binary mapping interpretations stated earlier provides the necessary foundation for the development of a probabilistic matching model. Since the similarity between two images are classified into two types based on the value of γ; similar images (γ = 1) and dissimilar images (0 ≤ γ < 1), then when matching two binary images we have,

$$\mathrm{P}(S, p) + \mathrm{P}(R, p) = 1 \qquad (6)$$

where P(*S*, *p*) and P(*R*, *p*) represent the probability of occurrence of an image pair being similar and dissimilar when matching two images, respectively. Furthermore, since dissimilar images are of two types based on the value of γ; quasi-dissimilar images (0 < γ < 1) and distinct-dissimilar images (γ = 0), thus,

$$\mathrm{P}(R, p) = \mathrm{P}(D, p) + \mathrm{P}(Q, p) \qquad (7)$$

where P(*D*, *p*) and P(*Q*, *p*) represent the probability of occurrence an image pair being distinct-dissimilar and quasi-dissimilar, respectively. However, if the pixel mappings are assumed to have equal probability of occurrences –as earlier stated, then this implies γ = 0 and P(*Q*, *p*) = 0, which implies that,

$$\mathrm{P}(R, p) = \mathrm{P}(D, p) \qquad (8)$$

Hence,

$$\mathrm{P}(S, p) + \mathrm{P}(D, p) = 1 \qquad (9)$$

Based on the mapping model and the image mapping variations it was shown that the probability of occurrence of distinct-dissimilar image pairs is equal to $1–2^{1-p}$, where *p* is the number of mappings.

Theorem: *Probabilistic Matching Model* (*PMM*)
Randomly mapping corresponding points between any two distinct-dissimilar (*D*) binary images (γ = 0) will ascertain the images to be dissimilar with probability P(*D*, p),

$$\mathrm{P}(D, p) = 1 - \frac{1}{2^{p-1}} \qquad p \ge 1 \qquad (10)$$

where *p* is the number of mappings.

□

Hence, P is the confidence level or *Detection Confidence* (*DC*) that an image pair is indeed distinct-dissimilar at a given mapping *p*. It is important to note that the theorem:

a) is only applicable to distinct-dissimilar images,
b) does not inform us anything about quasi-dissimilar or similar images, and
c) informs us of how fast distinct-dissimilar images can be found for a given confidence.

Furthermore, the theorem treats images as random processes and does not take into consideration any spatial relationship between image points. This negligence of image spatial data association might seem to be a weakness in the theorem at a first glance, but in reality it is an advantage and what gives this method its strength; it allows the process of detecting dissimilarity to be fast and not hindered by additional data processing of the images to understand spatial relationships among the pixels. This is similar to auto-correlation as a method used for template matching which ignores spatial relationships between pixels.

Examining (10), we see that P(*D*, *p*) has the following progression,

$$\mathrm{P}(D, p) = 0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \frac{31}{32}, \frac{63}{64}, \frac{127}{128}, \cdots \qquad (11)$$

A plot of P(*D*, *p*) is shown in Fig. 1 and the first 16 values are tabulated in Table II. It can be seen that P(*D*, *p*) quickly approaches unity with increasing value of *p* and in the limit as $p \to \infty$, we have,

$$\lim_{p\to\infty} \mathrm{P}(D, p) = \lim_{p\to\infty}\left(1 - \frac{1}{2^{p-1}}\right) = 1 \qquad (12)$$

Hence, for two completely distinct-dissimilar images, randomly mapping two points between them will ascertain the images to be dissimilar with 50% confidence. Mapping three points will ascertain the images to be dissimilar with 75% confidence, and mapping four points will ascertain the images to be dissimilar with 87.5% confidence. Confidence increases with increasing mapping; 93.75% confidence can be achieved by mapping five points, 96.875% confidence with 6 points, 99.219% confidence with 8 points, 99.902% with 11 points, 99.994% with 15 points and so on as dictated by (10). This mapping behavior agrees with our intuition that as we map more and more pixels between two different images, random pixel mapping will dominate.

Corollary 1
Given any two binary images, the mean number of random mappings required to ascertain that the two images are distinct-dissimilar with probability P is,

$$p(D,\mathrm{P}) = 1 - \log_2(1 - \mathrm{P}) \qquad 0 \ge \mathrm{P} \ge 1 \qquad (13)$$

□

Corollary 1 just states the inverse of (10), i.e. for a desired confidence level the minimum number of mappings can be computed.

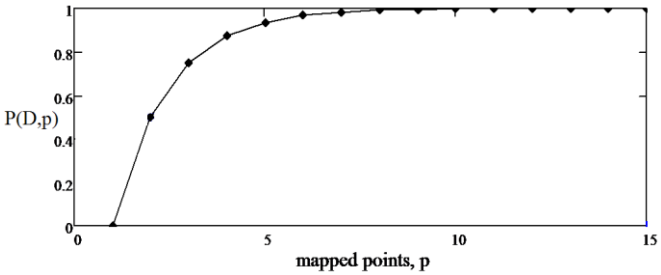

Fig. 1. Plot of P(*D*, *p*).

TABLE II
FIRST 16 VALUES OF P(*D*,*P*)

p	P(*D*, *p*)	*p*	P(*D*, *p*)
1	0.00000	9	0.99609
2	0.50000	10	0.99805
3	0.75000	11	0.99902
4	0.87500	12	0.99951
5	0.93750	13	0.99976
6	0.96875	14	0.99988
7	0.98438	15	0.99994
8	0.99219	16	0.99997

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

Corollary 2
For any image pair with similarity γ,

$$\lim_{\gamma \to 0} \mathrm{P}(Q = \gamma, p) = \mathrm{P}(D, p) \qquad (14)$$

□

This is due to the fact that as γ becomes smaller, the difference between the probabilities of occurrence of pixel mappings become smaller and in the limit when γ approaches zero, the image pair becomes distinct-dissimilar.

Corollary 3
For any probability P, the mean number of random mappings, *p*, required to ascertain that any two binary images are quasi-dissimilar with similarity γ,

1. cannot be less than the mean number of mappings required to ascertain that any two binary images are distinct-dissimilar,

$$p(Q = \gamma, \mathrm{P}) \geq p(D, \mathrm{P}) \qquad 0 \geq \mathrm{P} \geq 1 \qquad (15)$$

2. The difference between $p(Q = \gamma, \mathrm{P})$ and $p(D, \mathrm{P})$ is dependent on γ, and

$$p(Q = \gamma_1, \mathrm{P}) \geq p(Q = \gamma_2, \mathrm{P}) \quad \text{for} \quad \gamma_1 \geq \gamma_2 \qquad (16)$$

□

Hence, quasi-dissimilar image pairs with γ close to zero will have mapping-confidence correspondence values close to those for distinct-dissimilar image pairs. As γ increases and images become more similar, mapping values for given confidences will be higher than those for distinct-dissimilar image pairs. How the mapping value increases as a function of γ is yet to be determined, and is a topic of ongoing research.

The *PMM* theorem and its corollaries inform us of how quickly distinct-dissimilar image pairs can be detected for a given confidence level, but does not tell us how. However, based on our earlier discussion on mapping variations this can easily be accomplished. To detect dissimilarity when matching two images, random image points are repeatedly mapped between the images and the mapping tuple size is examined after every mapping:

- if the tuple mapping size is 1, then the images are *similar* and mapping is continued.
- if the tuple mapping size is 2 and the mapping variation is μ_5, μ_6, μ_9 or μ_{10}, then the images are *dissimilar*, and no further mapping is required.
- if the tuple mapping size is 3 then the image pair is *dissimilar* and no further mapping is required.

The *Mapping Rejection Number* (*MRN*), which is the number of mappings required to reject or detect a pair of images as being dissimilar, will be used to measure the performance of *PMM*. Furthermore, notation such as MRN_{DC} is frequently used as shorthand to denote the mapping rejection number for a given detection confidence.

VI. Discussion

In this section we present experimental results conducted on images and compare them to the theory developed. Tests with real images are presented.

A. Real Image Set #1

Fig. 2 shows the real images used in our work which are 128×128. The set consists of the following labeled images: 1) *urban*, 2) *mill*, 3) *forest1*, 4) *micro*, 5) *forest2*, 6) *trees*, 7) *phone*, 8) *coins*, 9) *cars*, 10) *appliances*, 11) *daisy* and 12) *park*. In our discussion below we will refer to an image by its label. However, when the images are tabulated or plotted as a group we will refer to the images by the indices indicated above.

Table III displays γ values for each image pair of the set. The image set has a similarity range of $0.002 \leq \gamma \leq 0.669$, with mean and standard deviation of 0.231 and 0.161, respectively. Two image pairs were found to be distinct-dissimilar ($\gamma < 0.01$): 1) the *urban-parking* image pair ($\gamma = 0.002$) and 2) the *mill-phone* image pair ($\gamma = 0.005$).
All images in the set were matched against each other producing 66 image pairs for a total of 66,000 trials (1000 trials per image pair). Table IV shows the results for $MRN_{0.999}$. The $MRN_{0.999}$ values have a range from 8 to 40 mappings, with a median value of 13 mappings, a mean value of 15.167 mappings and a standard deviation of 5.344. This implies that on average 99.9% of the images were found to be dissimilar by the 13th mapped value and for all images 99.9% of the trials were found to be dissimilar by the 40th mapped value. The theoretical value for distinct-dissimilar images is at a value of 10.966 mappings producing a set average discrepancy value of about 2 mappings and a set total discrepancy value of about 27 mappings. The discrepancy is attributed –as before- to the fact that the set γ values are not zero. The results of $MRN_{0.50}$, $MRN_{0.90}$, $MRN_{0.99}$, $MRN_{0.999}$ and $MRN_{0.9999}$ are summarized in Table V. Fig. 3 shows plots of the mapping results for all image pairs which clearly shows the increase of *MRN* values as γ increases. From these results we see,

- A small discrepancy between theoretical and experimental values due to the fact that image pairs are not distinct-dissimilar
- *MRN* increases with increasing γ value and is non-linear.
- *MRN* increases as the *DC* value increases.

Fig. 2. Test images. Left to right and from top to bottom: First row: *urban*, *mill*, *forest1* and *micro*. Second row: *forest2*, *trees*, *phone* and *coins*. Last row: *cars*, *appliances*, *daisy* and *park*.

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

B. *Image Set #2*

In this section we apply *PMM* to larger images to show that it is applicable to any image size with the same quick detection response. Fig. 4 shows the second real image set used in our work; four 3 mega-pixel binary images of size 3264×2448. The set consists of the following four labeled images: 1) *Tropical*, 2) *Seashore*, 3) *Children* and 4) *Kuwait_Skyline*. Similar to what was done before, the mean *MRN* values for *DC* = 0.50, 0.90, 0.99, 0.999 were obtained for each pair of images for 1000 trials. The corresponding *MRN* values for each *DC* value as well as the γ values for each image pair are tabulated in Table VI. The γ values for this set are in the range [0.121, 0.467] as shown in the table. From the table we see,

- as the similarity between images increases, *MRN* values increase.
- as the *DC* value increases the *MRN* increases for a given image pair.

This is consistent with the results obtained earlier for set #1. Furthermore, Table VII displays *MRN* results of this set with image pairs from set #1 of comparable γ values. For example, the image pair consisting of images #1 and #2 of this set has γ = 0.122 which is comparable to the image pair consisting of images #6 and #11 of set #1 (γ = 0.121), and they only differ by |Δγ| = 0.001. We see that the average *MRN* value difference between the two pairs is ε_{MRN} = 0.75 (less than a single mapping value). ε_{MRN} for comparable γ image pairs of the two sets varied from 0 to 2 mappings with an overall average value of 0.958. This discrepancy is an acceptable value and is attributed to the fact that the *MRN* values are rounded to the nearest integer value. These values are highly correlated producing a nearly perfect correlation value of 0.997.

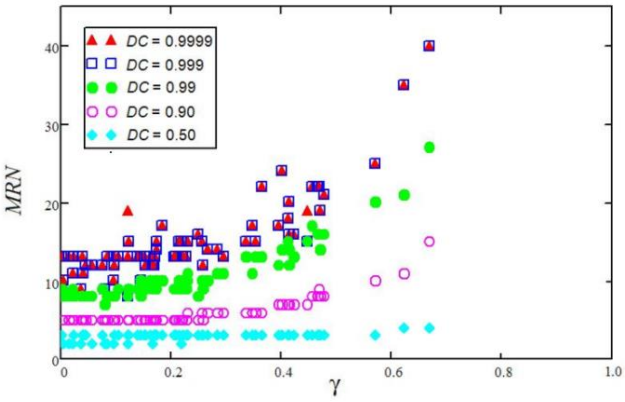

Fig. 3. *MRN* vs. γ for the real image set.

Fig. 4. Image Set #2: four 3Mega-pixel images.

TABLE III γ VALUES FOR REAL IMAGES SET #1

Image	1	2	3	4	5	6	7	8	9	10	11	12
1	–	0.255	0.220	0.022	0.040	0.145	0.016	0.229	0.228	0.168	0.173	0.002
2	0.255	–	0.447	0.122	0.222	0.424	0.005	0.477	0.456	0.402	0.205	0.023
3	0.220	0.447	–	0.077	0.415	0.364	0.145	0.669	0.571	0.467	0.259	0.056
4	0.022	0.122	0.077	–	0.155	0.046	0.267	0.096	0.042	0.036	0.123	0.139
5	0.040	0.222	0.415	0.155	–	0.229	0.354	0.395	0.347	0.250	0.172	0.163
6	0.145	0.424	0.364	0.046	0.229	–	0.085	0.413	0.414	0.337	0.121	0.097
7	0.016	0.005	0.145	0.267	0.354	0.085	–	0.081	0.104	0.097	0.185	0.208
8	0.229	0.477	0.669	0.096	0.395	0.413	0.081	–	0.623	0.469	0.284	0.174
9	0.228	0.456	0.571	0.042	0.347	0.414	0.104	0.623	–	0.471	0.296	0.151
10	0.168	0.402	0.467	0.036	0.250	0.337	0.097	0.469	0.471	–	0.214	0.154
11	0.173	0.205	0.259	0.123	0.172	0.121	0.185	0.284	0.296	0.214	–	0.169
12	0.002	0.023	0.056	0.139	0.163	0.097	0.208	0.174	0.151	0.154	0.169	–

TABLE IV $MRN_{0.999}$ FOR REAL IMAGE **SET #1**

Image	1	2	3	4	5	6	7	8	9	10	11	12
1	–	15	15	11	13	13	13	15	13	12	14	13
2	15	–	15	8	13	16	10	21	22	24	13	13
3	15	15	–	12	20	22	10	40	25	22	12	12
4	11	8	12	–	12	12	14	10	11	9	15	13
5	13	13	20	12	–	15	15	17	17	16	13	13
6	13	16	22	12	15	–	13	18	18	21	13	13
7	13	10	10	14	15	13	–	9	13	12	17	13
8	15	21	40	10	17	18	9	–	35	22	22	15
9	13	22	25	11	17	18	13	35	–	19	13	13
10	12	24	22	9	16	21	12	22	19	–	15	12
11	14	13	12	15	13	13	17	22	13	15	–	12
12	13	13	12	13	13	13	13	15	13	12	12	–

TABLE V $MRN_{0.999}$ FOR REAL IMAGE SET #1

DC	*Min*	*Max*	*Med*	μ	σ	*q*
0.50	2	4	3	2.86	0.42	6.78
0.90	5	15	5	5.92	1.71	3.45
0.99	7	27	9	10.80	3.68	2.93
0.999	8	40	13	15.16	5.34	2.83
0.9999	9	40	13	15.39	5.30	2.90

TABLE VI MRN_{DC} VALUES FOR IMAGE PAIRS OF SET #2

DC	1 & 2	1 & 3	1 & 4	2 & 3	2 & 4	3 & 4
0.50	4	4	4	3	4	4
0.90	6	8	8	6	6	9
0.99	10	14	15	11	10	15
0.999	13	16	21	13	13	21

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

VII. Conclusion

In this paper we have presented a probabilistic matching model for fast detection of dissimilar binary images. The method is based on a probabilistic model of the occurrence of image mappings. The method requires the comparison of only a few image points depending on the level of similarity between the images. If the images have no similarity and are distinct-dissimilar then there is a 50% confidence that the images will be found different by the 5th mapped pixel. By taking more pixels the confidence increases exponentially; a 75% confidence that dissimilar images will be detected by the 3rd mapping, a 93.75% confidence that dissimilar images will be detected by the 5th mapping, a 99.22% confidence that dissimilar images will be detected by the 8th mapping, a 99.90% confidence that dissimilar images will be detected by the 11th mapping and a 99.99% confidence that dissimilar images will be detected by the 15th mapping.

If the images have some similarity then more points need to be mapped and checked. Using the binary similarity measure γ as a reference for similarity measurement, image pairs with low similarity are detected at mapping values very close to the ideal case. On the other hand, image pairs with higher similarity required more points to attain the same confidence level. Images with good similarity ($\gamma \approx 0.7$), required as much as 40 points to be checked for a 99.99% confidence level.

An important aspect of our method that it is not image size dependent. This was proven by conducting tests on images of different sizes; 3 mega-pixel images showed results comparable to results obtained with 16k images, in agreement with the theory presented. In contrast, the majority of other methods used to detect dissimilar images are size dependent. Hence, the amount of time required for detection of dissimilar images by our method is only a fraction of the time required by other methods, as our preliminary additional research has shown and will soon report. Future work will concentrate on several other issues raised by this research, 1) developing a probabilistic model for quasi-dissimilar images, 2) template matching using *PMM*, and 3) a noise analysis of *PMM*.

References

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

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

[3] Adnan A. Mustafa, "A Framework for Quick Rejection of Dissimilar Binary Images", *International Journal of Signal Processing Systems*, Vol. 1, No.2, pp. 237-243, Dec. 2013. doi: 10.12720/ijsps.1.2.237-243.

[4] F. W. Leberl, *Radargrammetric Image Processing*, Artech House, Massachusetts, 1990.

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

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

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

[8] E. Baudrier, G. Millon, F. Nicolier and S. Ruan, "A fast binary-image comparison method with local-dissimilarity quantification", in the *18th International Conference Pattern Recognition*, pp. 216-219, 2006.

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

[10] F. Tang and H. Tao, "Fast multi-scale template matching using binary features", in the *8th IEEE Workshop on Applications of Computer Vision*, pp.36-39, 2007.

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

[12] A. Sleit, H. Saadeh, I. Al-Dhamari and A. Tareef, "An Enhanced Sub image Matching Algorithm for Binary Images", in the *American conference on Applied Mathematics*, pp.565-569. 2010.

[13] S. Choi, S. Cha and C. C. Tappert, "A Survey of Binary Similarity and Distance Measures", in *Journal of Systemics, Cybernetics and Informatics*, Vol 8 No 1 2010, pp 43-48.

[14] Adnan A. Mustafa, "A Modified Hamming Distance Similarity Measure for Quick Rejection of Dissimilar Binary Images", in the *International Conference on Computer Vision and Image Analysis*, pp. 180-185, 2015.

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

TABLE VII COMPARISON OF MRN_{DC} VALUES FOR IMAGE PAIRS OF COMPARABLE γ BETWEEN SETS #1 AND #2

	Set #2	**Set #1**	**Set #2**	**Set #1**	**Set #2**	**Set #1**	**Set #2**	**Set #1**	**Set #2**	**Set #1**	**Set #2**	**Set #1**
DC	**1 & 2**	**6 & 11**	**1 & 3**	**3 & 6**	**1 & 4**	**6 & 9**	**2 & 3**	**9 & 11**	**2 & 4**	**2 & 5**	**3 & 4**	**2 & 8**
0.50	4	3	4	3	4	3	3	3	4	3	4	3
0.90	6	5	8	6	8	7	6	6	6	5	9	8
0.99	10	9	14	13	15	12	11	11	10	10	15	16
0.999	13	13	16	19	21	18	13	13	13	13	21	21
γ	0.122	0.121	0.363	0.363	0.416	0.414	0.182	0.184	0.221	0.222	0.479	0.467
$\lvert\Delta\gamma\rvert$	0.001		0.000		0.002		0.002		0.001		0.012	
ε_{MRN}	0.75		1.75		2.00		0.000		0.5		0.75	

ISBN: 978-988-19253-2-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)