30-10-2004

Weighted segmented digital watermarking

Glen Wheeler
University of Wollongong, glenw@uow.edu.au

R. Safavi-Naini
University of Wollongong, rei@uow.edu.au

N. P. Sheppard
University of Wollongong, nps@uow.edu.au

Weighted segmented digital watermarking

Abstract

We introduce the notion of weighted watermarking for proof-of-ownership watermark protection of multimedia works that are the product of more than one author and where each author is considered to be of different importance relative to the other authors. We specifically examine weighted segmented watermarking for still images and generalise previous work on performance measurement of watermark embedding patterns in the presence of cropping attacks.

Keywords

digital watermarking

Disciplines

Physical Sciences and Mathematics

Publication Details

This paper was originally published as: Wheeler, GE, Sheppard, NP & Safavi-Naini, R, Weighted segmented digital watermarking, Third International Workshop on Digital Watermarking 2004 (IWDW 2004), Seoul, South Korea, October 30-November 1 2004, 89-100.

Weighted Segmented Digital Watermarking

Glen E. Wheeler, Reihaneh Safavi-Naini, and Nicholas Paul Sheppard

School of Information Technology and Computer Science
The University of Wollongong NSW 2522
Australia
{gew75,rei,nps}@uow.edu.au

Abstract. We introduce the notion of weighted watermarking for proof-of-ownership watermark protection of multimedia works that are the product of more than one author and where each author is considered to be of different importance relative to the other authors. We specifically examine weighted segmented watermarking for still images and generalise previous work on performance measurement of watermark embedding patterns in the presence of cropping attacks.

1 Introduction

Many multimedia works are the product of more than one author, and it is often the case that the relative importance of one contribution to a work is greater than that of another contribution. In this paper, we consider proof-of-ownership watermarking in which it is desirable for this disparity in importance to be reflected in watermark protection of the work.

More specifically, where one contributor is considered to be more important than others, we would like to afford greater protection to that contributor than to less important contributors. That is, it should be harder for an attacker to successfully remove that contributor's watermark than other watermarks.

Multiple watermarking, and, in particular, segmented watermarking [4] is one method by which the contribution of several authors to a work can be recognised by a multimedia protection system. In this paper, we introduce the notion of *weighted* segmented watermarking in which each contributor to a multimedia work is assigned an integer weight reflecting his or her relative importance amongst the contributors to the work, and contributors are afforded greater or lesser levels of protection according to their assigned weight.

In segmented watermarking, the work to be watermarked is divided into a series of individual *segments* and each segment is watermarked independently. In particular, in this paper we consider segmented watermarking of still images in which segments are formed by dividing the image into square blocks, each of which contains one contributor's watermark. If a watermark is present in one or more segments of the work, the owner of that watermark is reported to be an owner of the work as a whole by an arbiter.

An obvious attack on this kind of system is to crop the image so that one or more of the watermarked segments is (either intentionally or coincidentally)

removed from the work. In general, each contributor to a work may have his or her watermark contained in more than one segment, which gives some resistance to a cropping attack, but if all of the segments assigned to one contributor are removed, the attack can be considered successful.

We will generalise the work of Atallah and Frikken [1] and Scealy, et al. [3] to examine watermark embedding patterns that minimise the risk of watermarks being destroyed in this way, while taking into account the weighting assigned to contributors. We present generalised metrics for comparing the performance of weighted embedding patterns in the face of cropping attacks, and compare the performance of several embedding patterns using the generalised metrics.

2 Previous Work

Guo and Georganas [2] propose a "joint watermarking" scheme based on cryptographic secret sharing techniques. In their scheme, part of the watermark pattern is made to be dependent on a shared secret which can only be recovered if some approved subset of the watermark owners comes together and re-constructs the secret according to the underlying secret sharing scheme.

Many secret-sharing schemes have mechanisms by which some parties can be made more important than others in the sense that fewer parties with important shares are required to re-construct the secret than parties with less important shares. In this way, some watermark owners can be made more important than others by distributing more important shares to these owners.

Guo and Georganas' approach gives an all-or-nothing result: either a given set of watermark owners can recover the watermark completely, or they cannot recover it at all. In weighted watermarking, watermark detection degrades gracefully when one or more watermark owners are unavailable, and, furthermore, the detector can recover the relative importance of each watermark owner by comparing the detector response for each owner.

Atallah and Frikken [1] and Scealy et al. [3] describe some performance metrics for cropping-resistance of square-block segmented watermarking of still images, as we do in this paper. These metrics are based on the notion of *completeness*: a region of an image is said to be *complete* if and only if it contains at least one copy of every watermark from each contributor. In these papers, all watermark owners were considered to be equally important.

Atallah and Frikken define a worst-case metric called the *maximum non-complete area* (MNCA) as the largest rectangular region of a watermarked image missing at least one watermark (i.e. is not complete), which gives an indication of the largest possible cropping attack that might succeed. We will later define a generalised form of this metric to account for the existence of watermark owners of disparate importances.

Scealy, et al. further define two average-case metrics called the *all-exclusion* and *single-exclusion* metrics. The all-exclusion metric gives a measure of how likely an arbitrary cropping attack is to succeed in removing a watermark owner;

the single-exclusion metric is similar but applies greater penalties if cropping attacks can remove multiple watermark owners. In this paper, the degree of success of a cropping attack is measured by the relative importance of the watermark owners that are removed, and we will define a generalised form of the all-exclusion metric in order to capture this.

3 Definitions

In this paper, we consider watermarking of a rectangular still image X by dividing it into a series of $t \times t$ square blocks $X(1,1), X(1,2), \ldots$. Each block $X(x,y)$ contains one watermark according to some underlying still-image watermarking scheme. We have m watermark owners $1, \ldots, m$ and each author i has

- some watermark q_i according to the underlying watermarking scheme; and
- an integer weight w_i measuring his or her relative importance amongst the authors, with higher weights indicating greater importance.

We assume that the weights are assigned by some dealer and that watermarking is done by some embedder under the control of the dealer. The protection system should ensure that protection is distributed according to the supplied weights, i.e. authors with higher weights should be better protected in some sense than authors will lower weights.

An *embedding pattern* is a mapping $\phi : \mathcal{N} \times \mathcal{N} \to \{1, \ldots, m\}$ mapping an image segment $X(x,y)$ to a watermark owner $\phi(x,y)$. The design and evaluation of embedding patterns that reflect the weighting assigned to the owners is the subject of this paper.

We will use the same model for a cropping attack as that used by Scealy, et al., that is, as the selection of a rectangular area of the segementation grid. We can assume this because

- non-rectangular images are unlikely to have any value for proving ownership due to the obvious nature of the tampering; and
- each partial segment included in a cropped image is either large enough for reliable watermark detection (in which case this segment can be considered wholly present), or it is not.

4 Metrics

The desirability of a given embedding pattern for a given set of weights can be judged according a variety of different measurements reflecting different aspects of the pattern. In this section, we propose two metrics that measure the worst-case and average-case performance of the pattern.

4.1 Generalised Maximum Non-Complete Area (GMNCA)

Atallah and Frikken [1] define an area of the segmentation grid to be *complete* if and only if at least one copy of each watermark is present in that area. The *maximum non-complete area* (MNCA) of an embedding pattern ϕ is then the size of the largest rectangle of ϕ in which at least one watermark is not present. The maximum non-complete area is a measure of the largest possible region that an attacker can crop from the image while removing at least one watermark; all larger regions contain all watermarks.

We define the *generalised maximum non-complete area* as a vector $\alpha = (\alpha_1, \ldots, \alpha_m)$ with α_i being the size of the largest rectangle of ϕ not containing at least one copy of watermark q_i.

4.2 Generalised All-Exclusion (GAE)

An embedding pattern is said to be *periodic* if $\phi(\delta x, \delta y) = \phi(x, y)$ for some fixed integers δx and δy. Scealy, et al. define the "all-exclusion metric" in terms of the proportion of *minimal cropping regions* found to be complete over one period of the embedding pattern. All of the embedding patterns used by Scealy, et al., and also in this paper, are periodic.

Scealy, et al. determine the set of all minimal cropping regions for an area T as the set of all rectangles with area at least T, and minimal in each dimension. Formally, a minimal cropping region for area T is an $a \times b$ rectangle such that

- if $a \leq b$ and $a \leq \sqrt{T}$, then $b = \lceil \frac{T}{a} \rceil$; and
- if $a > b$ and $b \leq \sqrt{T}$, then $a = \lceil \frac{T}{b} \rceil$.

For example, the minimal cropping regions for $T = 5$ are the rectangles of size 5×1, 3×2, 2×3 and 1×5.

Scealy, et al. argue that testing all minimal cropping regions for area m (the number of watermark owners) gives a good indication of how the embedding pattern can be expected to perform in the face of an arbitrary region being cropped from the image. This is because all regions smaller than the minimal cropping regions for m cannot possibly be complete, while all regions larger than the minimal cropping regions for m encompass one or more minimal cropping regions. Thus the all-exclusion metric gives an indication of the likelihood of success for an arbitrary cropping attack, while reducing the complexity of the test as compared to testing all possible cropping regions.

Let $C = \{C_1, \ldots C_r\}$ be the set of all minimal cropping regions for area m over one period of the embedding pattern ϕ. Note that the embedding pattern is considered to "wrap around" at the edges so that minimal cropping regions at the edges of the period will effectively extend into the next period of the pattern. We define the *generalised all-exclusion* (GAE) metric to be a vector $\eta = \eta_1, \ldots, \eta_m$ with

$$\eta_i = \frac{|\{C_k \in C : q_i \notin C_k\}|}{|C_k|}, \tag{1}$$

that is, the proportion of minimal cropping regions that do not contain watermark q_i. The higher the value of η_i, the more likely q_i is to be eliminated by a cropping attack.

Note that our definition for generalised all-exclusion is the dual of that used for ordinary all-exclusion by Scealy, et al., in which all-exclusion is defined as the proportion of minimal cropping regions that are complete, that is, *do* contain the required watermarks. By using the dual definition, all of the metrics considered in this paper indicate better performance by lower values.

5 Evaluation of Embedding Patterns

5.1 Mean

In general, lower values for the elements α_i and η_i indicate that the corresponding watermark q_i is more difficult to remove by cropping. An embedding pattern that reduces these values therefore reduces the susceptibility of a watermarked image to cropping attacks.

A simple method of measuring the overall resistance of the embedding pattern to cropping attacks is to take the means of the GMNCA and GAE vectors. We will therefore define the *GMNCA mean* to be

$$\mu_\alpha = \frac{\sum_{i=1}^{m} \alpha_i}{m} \tag{2}$$

and the *GAE mean* similarly as

$$\mu_\eta = \frac{\sum_{i=1}^{m} \eta_i}{m} \tag{3}$$

Lower values of μ_α and μ_η indicate greater overall resistance to cropping attacks.

5.2 Divergence

If owner i has a high weight w_i, then a good embedding pattern should have

- only relatively small areas not containing q_i (i.e. low α_i); and
- relatively few minimal cropping regions from which q_i is absent (i.e. low η_i).

We can quantify this in terms of the products $w_i\alpha_i$ and $w_i\eta_i$: we would like $w_i\alpha_i$ to be roughly the same for all i, and similarly for $w_i\eta_i$.

We define the *GMNCA divergence* θ_A to be the angle between the GMNCA product vector $\hat{\alpha} = (w_1\alpha_1, \ldots, w_m\alpha_m)$ the all one vector, that is

$$\cos(\theta_\alpha) = \frac{\hat{\alpha} \cdot \mathbf{1}}{|\hat{\alpha}|\sqrt{m}} \tag{4}$$

where $\mathbf{1} = (1, \ldots, 1)$. A GMNCA divergence close to zero indicates that the embedding pattern is more faithful to the supplied weights insofar as it is harder to find regions without q_i if w_i is high than if w_i is low.

We similarly define the *GAE divergence* θ_η to be the angle between the GAE product vector $\hat{\eta} = (w_1\eta_1, \ldots, w_m\eta_m)$ and the all one vector, that is,

$$\cos(\theta_\eta) = \frac{\hat{\eta} \cdot \mathbf{1}}{|\hat{\eta}|\sqrt{m}} \tag{5}$$

As for the GMNCA divergence, a GAE divergence of close to zero indicates that the embedding pattern is more faithful to the supplied weights.

6 Embedding Patterns

An obvious method of distributing watermarks according to their relative weights is to form a basic pattern in which each watermark q_i appears w_i times, and then repeat this pattern as necessary to fill the image to be watermarked. All of the embedding patterns considered in this paper follow this paradigm, and are differentiated in the way they determine the basic pattern and in the way they repeat the pattern throughout the image.

6.1 Cyclic Embedding

Scealy, et al. [3] show that their metrics favour embedding patterns based on the *cyclic* paradigm, in which each row of the embedding pattern is a cyclic shift of the row above it (and similarly for columns).

Given a set of watermarks and corresponding weights, we can form a vector S of length $\ell = \sum_{i=1}^{m} w_i$ with w_i elements set to q_i and use this vector as the input to the cyclic embedding algorithm. In this way, watermarks will appear with frequency proportional to their weight.

Given an initial vector S, a cyclic embedding can be defined as

$$\phi(x, y) = S((xH + yJ \bmod m) + 1) \tag{6}$$

for some integer step sizes H and J. Scealy, et al. set $J = 1$ for all of their experiments, and observe that H is usually best chosen to be as large as possible – specifically, equal to the largest number less than $\lceil \frac{\ell}{2} \rceil$ that is relatively prime to ℓ – in order to maximise the difference between two adjacent rows of the embedding.

It remains to determine how the initial vector S should be arranged. Intuitively, for resistance to cropping, we want each watermark to be spread evenly throughout the image, since clustering a watermark in one area will lead to large areas elsewhere in which it is not present. Without loss of generality, assume that the weights $w_1, \ldots, w_m$ are arranged in non-increasing order. Then the following is a simple algorithm for distributing the watermarks $q_1, \ldots, q_k$ evenly through a string such that each watermark appears a number of times equal to its weight:

create w_1 strings $S_1, \ldots, S_{w_1}$ with $S_i = q_1$
set $k = 1$

for $i = 2$ to m
    for $j = 1$ to w_i
        append q_i to S_k
        set $k = k + 1$
    end for
end for
set $S = S_1 \parallel \cdots \parallel S_{w_1}$

where '$\parallel$' denotes concatenation. We will use this algorithm in all of our experiments.

Figure 1 shows a vector S and cyclic embedding pattern ϕ for four authors with $w = 3, 2, 2, 1$. The q's have been omitted from the figures for clarity, that is, only the authors' numbers are shown.

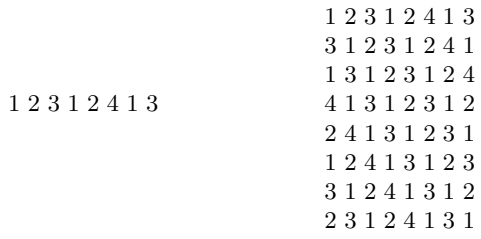


Fig. 1. (a) S and (b) one period of ϕ for cyclic embedding with $w = 3, 2, 2, 1$ and $H = J = 1$.

6.2 Tile Rotation

Scealy, et al. also propose a "tile rotation" method in which a basic pattern is rotated each time it is repeated, so that (if the image is large enough) every watermark appears at least once in every row and every column. This method obtains better results than the cyclic method for some special values of m where the cyclic method performed relatively poorly.

Let L be an $a \times b$ matrix such that

$$ab = \ell = \sum_{i=1}^{m} w_i \tag{7}$$

with exactly w_i entries set to q_i for all $1 \le i \le m$. Scealy, et al. then derive the formula

$$\phi(x, y) = L(x + \lfloor \frac{x}{a} \rfloor \bmod a, y + \lfloor \frac{y}{b} \rfloor \bmod b) \tag{8}$$

for rotating the basic tile L over the whole image.

If $a = 1$ or $b = 1$, this method reduces to the cyclic method since L becomes a vector, and this vector is rotated for every row of the embedding pattern. Scealy, et al. do not allow $a = 1$ or $b = 1$, but their results show that the tile rotation method consequently performs poorly for prime m (which is equal to ℓ when there are no weights), since in this case L has empty entries. In this paper, we will use $a = \ell$ and $b = 1$ for prime ℓ so that this method will perform as well as the cyclic method in these cases.

In choosing the layout of the base tile L, we have a similar problem to the one we had in choosing the layout of the initial vector S for cyclic embedding. We can define the initial tile by use of the algorithm described in Section 6.1: given the output vector S, we place the first a elements of S on the first row of L, the second a elements on the second row, and so on, that is

$$L(x, y) = S(a(y - 1) + x). \tag{9}$$

Figure 2 shows the tile rotation pattern for the same parameters used in Fig. 1, with $a = 4$ and $b = 2$. Again, the q's have been omitted for clarity.

	1 2 3 1 2 4 1 3
	2 4 1 3 1 2 3 1
	1 1 2 3 3 2 4 1
1 2 3 1	3 2 4 1 1 1 2 3
2 4 1 3	3 1 1 2 1 3 2 4
	1 3 2 4 3 1 1 2
	2 3 1 1 4 1 3 2
	4 1 3 2 2 3 1 1
(a)	(b)

Fig. 2. (a) S and (b) one period of ϕ for tile rotation embedding with $w = 3, 2, 2, 1$ and $a = 4$ and $b = 2$.

7 Experiments

In general, we would expect different weight vectors to result in different performances from any given embedding method. There is an infinite number of possible weight vectors, though not all of them seem very likely in practice. We ran two series of tests:

- the *lead authorship* case in which a single important author is assigned a weight of 2 and all other authors are assigned a weight of 1;
- the *bimodal authorship* case in which the authors are evenly divided into two groups, with the more important group being assigned weight 2 and the

less important group being assigned weight 1. For odd m, there is one more author in the less important group than in the more important group.

Lead authorship models the case where artistic direction is taken by a single lead author (or production company), who is then assisted by secondary authors to fill out details. Bimodal authorship of tests models a collaboration in which a core group of designers outsource minor tasks to other authors.

For each set of weights, we computed the mean and divergence of the GMNCA and GAE for each of the embedding paradigms described in Section 6 for 2 up to 10 authors. For the cyclic method, we chose $J = 1$ and H to be the largest integer less than $\lceil \frac{\ell}{2} \rceil$ relatively prime to ℓ, as suggested by Scealy, et al.

Figure 4 shows the GMNCA mean and divergence for each embedding pattern and each set of weights, using the graph legend shown in Fig. 3. Figure 5 similarly shows the GAE mean and divergence.

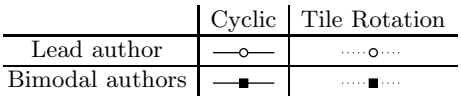


Fig. 3. Graph legend

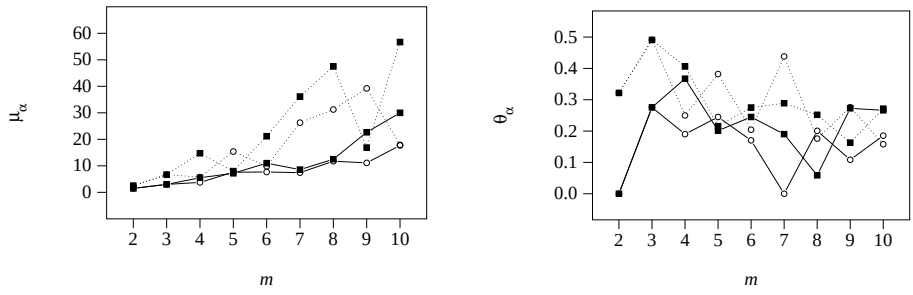
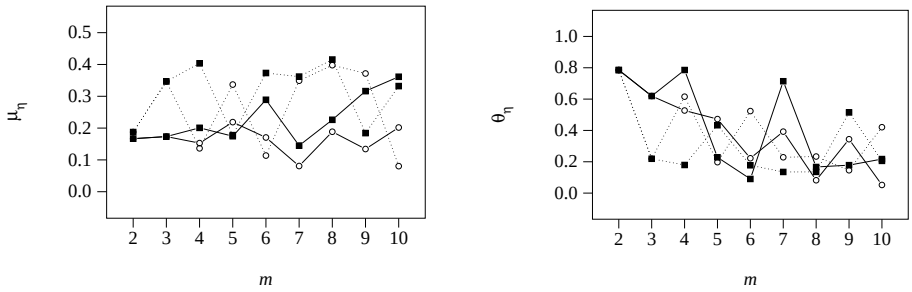


Fig. 4. (a) GMNCA mean and (b) GMNCA divergence

8 Discussion

The results are very similar for both models of authorship used for our experiments. The cyclic method obtains better results in most cases, though there are



(a) (b)

Fig. 5. (a) GAE mean and (b) GAE divergence

a significant number of cases in which the tile rotation method has scored better, particularly for the GAE metric. The cyclic method appears to be somewhat less erratic than the tile rotation method, however.

As we would expect, the GMNCA mean increases as the number of authors is increased, since more area is required to fit in more watermark owners. The GAE mean is much more erratic than the GMNCA mean and it is difficult to draw conclusions about any trends that the GAE mean may show, though there is at least arguably a slight upward trend as we would expect.

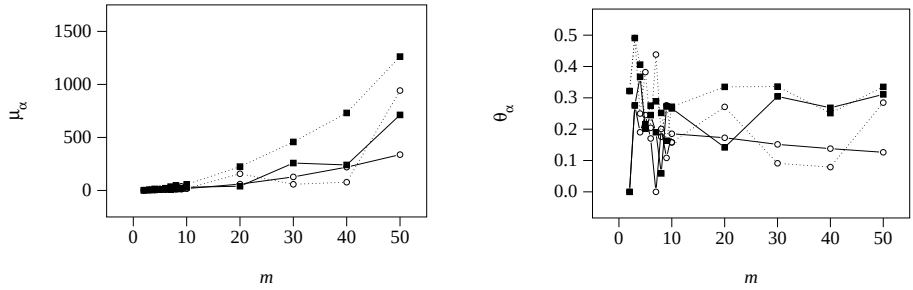
Unlike the means, the divergence measures seem to exhibit a downward trend as the number of authors is increased, though this is slight and rather erratic except in the GAE divergence of lead authorship (Fig. 5(b)). In the latter case, this might be explained by the embedding pattern becoming closer and closer to the simple case in which all authors have equal weight, and in which divergence is zero for any reasonable embedding pattern.

In the other divergence cases, the results may be too erratic to draw conclusions with great confidence, but it seems plausible to suggest that the behaviour of an embedding pattern is "evening out" as it grows larger. For a greater number of authors, the period of the embedding pattern and the number of minimal cropping regions increases, and a greater number of tests are performed. The sample population used to compute the divergence score is thus larger and may therefore show less variance.

In general, our generalised measures are somewhat more erratic than the unweighted measures reported by Scealy, et al. Scealy, et al. note that certain numerical properties of m – such as the whether or not there are numbers near $\frac{m}{2}$ relatively prime to m in the cyclic method, or m is prime in the tile rotation method – have a significant impact on the formation of embedding patterns. In introducing weights, we may have increased the opportunity for some chance property of the input parameters to dramatically affect the properties of the embedding pattern. For larger values of m we might expect the proportion of

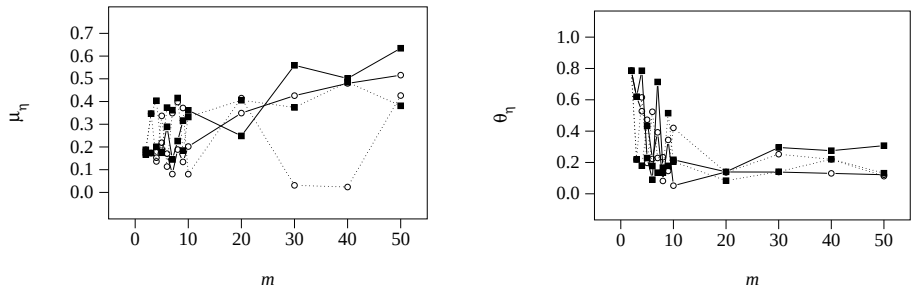
troublesome m's and w's to grow smaller, also contributing to a reduced variance in the results.

Figures 6 and 7 show graphs of the GMNCA and GAE metrics, respectively, for $m = 10, 20, \ldots, 50$ in addition to the smaller m's shown in the earlier graphs. We stopped at $m = 50$ as the amount of computation required to calculate the metrics becomes prohibitive for larger m. The extended graphs confirm that the means increase as m increases, as we would expect. The divergences appear to decrease slightly, but find a level at around $m = 20$, after which they do not decrease any further.



(a) (b)

Fig. 6. (a) GMNCA mean and (b) GMNCA divergence for large m



(a) (b)

Fig. 7. (a) GAE mean and (b) GAE divergence for large m

Of course, it does not seem very likely that a single image would be the result of the collaboration of an extremely large collection of authors and so the behaviour of the divergence (or any other measurement) for very high m may

not hold much interest in practice, at least in the image case we are considering here. For larger works such as video where large numbers of authors may be more realistic, our metrics would need to be extended to their three-dimensional forms.

9 Conclusion

We have introduced the notion of weighted segmented digital watermarking, and generalised previous work on cropping-resistance in segmented watermarking to provide performance measures for the weighted case. As in the unweighted case embedding patterns based on the cyclic paradigm typically give the best results though the similar tile rotation method sometimes obtains better scores.

The addition of weights, however, has made the results somewhat more erratic than observed in the unweighted case. We have conjectured that this is due to the greater number of interacting parameters used in forming embedding patterns. In small objects, such as still images, this erratic performance may be inevitable given the difficulty of satisfying a large number of parameters in a relatively small solution space.

Our metrics, and the earlier ones from which they have been derived, are quite narrow in that they measure only the effectiveness of an embedding pattern in defeating a cropping attack specific to segmented watermarking. For a complete comparison of multiple watermarking techniques, more broad-based metrics need to be defined. These broader metrics are likely to be computed in quite different ways to the metrics presented in this paper, though we think notions such as the division between mean (measuring overall goodness) and divergence (measuring faithfulness to a particular parameter set) may also be useful in a broader sense than the sense in which we have used them here.

10 Acknowledgements

This work was partly funded by the Co-operative Research Centre for Smart Internet Technology, Australia.

References

1. M. Atallah and K. Frikken. Cropping-resilient segmented multiple watermarking. In *Workshop on Algorithms and Discrete Structures*, pages 231–242, 2003.
2. H. Guo and N. D. Georganas. A novel approach to digital image watermarking based on a generalized secret sharing scheme. *Multimedia Systems Journal*, 9:249–260, 2003.
3. R. Scealy, R. Safavi-Naini, and N. P. Sheppard. Performance measurement of watermark embedding patterns. In *International Workshop on Digital Watermarking*, pages 77–85, Seoul, Korea, 2003.
4. N. P. Sheppard, R. Safavi-Naini, and P. Ogunbona. On multiple watermarking. In *Workshop on Security and Multimedia at ACM Multimedia*, pages 3–6, Ottawa, Canada, 2001.