EUROGRAPHICS 2000 / M. Gross and F.R.A. Hopgood
(Guest Editors)

Volume 19 (2000), Number 3

Robust Motion Watermarking based on Multiresolution Analysis

Tae-hoon Kim, Jehee Lee, and Sung yong Shin

Department of Computer Science, Korea Advanced Institute of Science and Technology,
373-1 Kusung-dong Yusung-gu, Taejon, Republic of Korea
e-mail: polphit@jupiter.kaist.ac.kr, jehee@jupiter.kaist.ac.kr, and syshin@cs.kaist.ac.kr

Abstract

Digital watermarking is one of commonly used solutions for copyright protection. A watermark should be imperceptible and robust to various attacks. In this paper, we address watermarking for motion data. Our watermarking scheme is based on two well-known ideas, so called multiresolution representation and spread spectrum. We embed a watermark into a motion signal by perturbing large detail coefficients of its multiresolution representation, and extract the watermark by analyzing perturbation of coefficients from a suspected signal. For more effective watermark extraction, we align suspected motion data to the original using dynamic time warping. Our scheme has merits of spread spectrum such as the resilience to common signal processing as well as the robustness to time warping.

1. Introduction

The advent of motion capture systems offers a convenient means of acquiring realistic motion data, that is, capturing live motion of a real actor. Due to the success of this technology, realistic and highly detailed motion data are rapidly spreading in computer animation. Archives of motion clips are also commercially available. The proliferation of motion data gives rise to demand on copyright protection for those data. One of widespread approaches for copyright protection is digital watermarking, that is, embedding an ownership information into data. For this purpose, the embedded information, called a watermark, is an imperceptible identification code that permanently remains in the data unless they are extremely degraded.

The process of publishing watermarked data and proving ownership claim is as follows: Suppose that one creates original data and embeds a watermark into them, and then publishes the watermarked data while keeping the original and the embedded watermark in secret. The published data may be altered by an attacker, and then published as if they were the attacker's original work. Finding such suspected data, the author compares it with the original to extract the watermark that may remain in the suspected. Finally, the author analyzes the similarity between inserted and extracted watermarks. If the similarity is high enough to claim an ownership, the author can prove that the suspected data are illegal copy of his (or her) work.

A common approach to embed a watermark is due to spread spectrum[31, 32], which is a mechanism transmitting a signal over a much larger bandwidth than normally required. Such an approach enjoys various benefits from nice properties such as jam resistance (JR), low probability of intercept (LPI), and so on. The property of JR provides a degree of resistance to interference and jamming. The LPI property provides a means of decreasing the probability of intercept by an adversary. Spread spectrum-based watermarking transforms a signal to that in the frequency domain, and then embeds a random watermark generated from a secret key. The watermark, embedded with this scheme, is undetectable and robust to various attacks due to the favorable properties of spread spectrum.

The motion data of an articulated figure consists of a bundle of motion signals. Every signal represents a sequence of sampled values each of which corresponds to either the position or the orientation of a part of the figure, called a link or segment. Unlike the position of a 3-dimensional object, its orientation cannot be expressed by a vector in a 3-dimensional space without yielding any singularity. The non-singular representations, such as rotation matrices and unit quaternions, form a *Lie group* which cause a complication in obtaining frequency information from the motion data. Thus, digital watermarking for motion data requires extra efforts for handling orientation data.

In this paper, we present a new scheme of embedding watermarks into a motion signal based on the multiresolution representation proposed by Lee[28]. This representation of a motion consists of a coarse base signal and a hierarchy of motion displacement maps. Displacement mapping is originally invented for warping a canned motion while preserving its detail characteristics[5, 44]. In the context of multiresolution analysis, the hierarchy of displacement maps is used for adding details sequentially to the base signal to reproduce the original motion encoded in the representation. The displacement maps are computed by two basic operations: reduction and expansion. The expansion is commonly achieved by a subdivision scheme that can be considered as up-sampling followed by smoothing. The reduction is a reverse operation, that is, smoothing followed by downsampling. The generalized version of interpolatory subdivision schemes and smoothing filters are used to process motion signals that include orientations as well as positions.

The remainder of the paper is organized as follows. After a brief review of watermarking and multiresolution representations, we present a multiresolution structure for representing a motion and describe in detail how to construct it in section 3. In section 4, we discuss a motion watermarking scheme based on the multiresolution representation. In section 5, we provide experimental results. Finally, we conclude this paper in section 6.

2. Related Work

In this section, we describe previous works about watermarking and multiresolution representations.

2.1. Watermarking Methods

Digital watermarking is one of commonly used solutions for copyright protection although it may also be used for other purposes such as data authentication, fingerprinting, and secret data hiding as well. Watermarking schemes are classified into forensic watermarking[7, 8, 34, 36, 37] and blind detection methods[20, 43]. The former requires original data, and the latter needs only a secret key for detection. The forensic watermarking is more robust than the blind detection method, while the blind detection is more suitable to product monitoring in network distribution or broadcasting[43]. In this paper, we focus on the forensic watermarking scheme for copyright protection.

There are several watermarking methods proposed for various media such as image, audio, video, and 3D geometric models. The early watermarking methods include modifying the least significant bit of the data[40], embedding a secret information that resembles quantization noiseinto a dithered image [42], and etc. The reader is referred to Bender *el al.*[1] and Cox *et al.*[8] for excellent surveys of the early approaches.

Watermarking schemes achieved more imperceptibility and robustness by transforming the data using multiresolution analysis[34], Fourier transform[36], or discrete cosine transform[7, 8]. These transform-domain-based approaches can embed a watermark into wide portion of the data without incurring noticeable visual artifacts. To prevent a watermark from being removed without extreme degradation of the data, the watermark is also embedded globally into the most perceptually significant portions of the data by analyzing frequency domain. Ruanaidh *et al.*[36, 37] modified the phase values of Fourier coefficients to convey information. Cox *et al.*[7, 8] used a spread-spectrum method for information embedding to achieve more robustness. Praun *et al.*[34] addressed robust 3D mesh watermarking by generalizing the spread-spectrum approach for arbitrary triangular meshes.

In the other schemes, the human visual system or the human auditory system was used to generate a more effective watermark. Boney *et al.*[4] generated a watermark by approximating the frequency masking characteristics of the human auditory system. Podilchuk and Zeng[33] employed visual models to determine image dependent upper bounds on watermark insertion, and increased robustness to common image modification.

2.2. Multiresolution Representations

Multiresolution representations are enormously popular in computer graphics applications such as curve and surface editing [6, 14, 16, 25, 30, 39], polygonal mesh editing [19, 24, 26, 45], image editing [2], image querying [23], texture analysis and synthesis [3, 21], video editing and viewing [15], image and surface compression [10, 11], global illumination [18], and variational modeling [17].

Multiresolution representations have been used for motion editing and synthesis as well. Liu *et al.*[29] employed hierarchical wavelets with adaptive refinement to provide a significant speedup for spacetime optimization. Bruderlin and Williams[5] used a digital filterbank technique to store motion data as a hierarchy of detail levels, where each level represents a different band of spatial frequencies. Through the use of the hierarchy, the motion data can be modified interactively by amplifying/attenuating particular frequency bands and a new motion can be generated by blending two existing motions band-wisely.

3. Multiresolution Analysis

We employ the multiresolution representation of Lee [28], that can be constructed through two basic operations: reduction and expansion (see Figure 1).

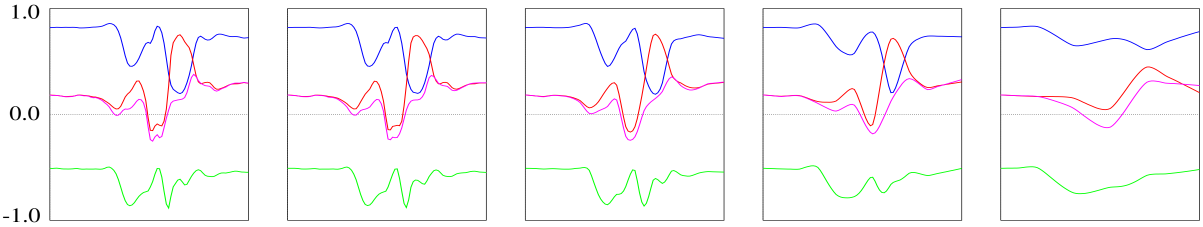

Figure 2: *Multiresolution analysis for a live-captured signal. The four curves represent the change of w-, x-, y-, and z-components, respectively, of a unit quaternion with respect to time. (from left to right) Original signal,* $\mathbf{M}^{(3)}, \mathbf{M}^{(2)}, \mathbf{M}^{(1)}, \mathbf{M}^{(0)}$

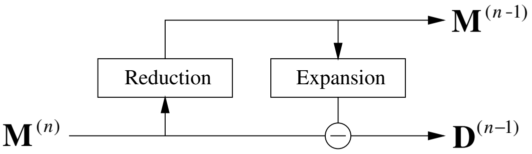

Figure 1: *Wiring diagram of the multiresolution analysis*

3.1. Displacement Mapping

The configuration of an articulated figure is specified by its joint configurations in addition to the position and orientation of the root segment. For uniformity, we assume that the configuration of each joint can be specified by a 3-dimensional rigid transformation. Then, we can describe the DOFs at every body segment as a pair of a vector in $\mathbb{R}^3$ and a unit quaternion in $\mathbb{S}^3$.

The *motion data* for an articulated figure comprise a bundle of *motion signals*. Each signal consists of a sequence of frames, $\{(\mathbf{p}_i, \mathbf{q}_i) \in \mathbb{R}^3 \times \mathbb{S}^3\}$, that correspond to the position and orientation of a body segment. Each frame $(\mathbf{p}_i, \mathbf{q}_i)$ specifies a rigid transformation $T_{(\mathbf{p}_i,\mathbf{q}_i)}$ that maps a point in $\mathbb{R}^3$ to another:

$$T_{(\mathbf{p}_i,\mathbf{q}_i)}(\mathbf{u}) = \mathbf{q}_i \mathbf{u} \mathbf{q}_i^{-1} + \mathbf{p}_i, \tag{1}$$

where $\mathbf{u} = (x,y,z) \in \mathbb{R}^3$ is considered as a purely imaginary quaternion $(0,x,y,z)$. Given two motion signals $\mathbf{M} = \{(\mathbf{p}_i, \mathbf{q}_i) \in \mathbb{R}^3 \times \mathbb{S}^3\}$ and $\mathbf{M}' = \{(\mathbf{p}'_i, \mathbf{q}'_i) \in \mathbb{R}^3 \times \mathbb{S}^3\}$, the motion displacement $\mathbf{D} = \{(\mathbf{u}_i, \mathbf{v}_i) \in \mathbb{R}^3 \times \mathbb{R}^3\}$ between them measured in a local (body-fixed) coordinate system gives rise to a transformation, $T_{(\mathbf{p}'_i,\mathbf{q}'_i)} = T_{(\mathbf{p}_i,\mathbf{q}_i)} \circ T_{(\mathbf{u}_i,\exp(\mathbf{v}_i))}$, called a motion displacement mapping. We compactly represent this equation by $\mathbf{M}' = \mathbf{M} \oplus \mathbf{D}$ or $\mathbf{D} = \mathbf{M}' \ominus \mathbf{M}$, where

$$\begin{aligned}(\mathbf{p}'_i, \mathbf{q}'_i) &= (\mathbf{p}_i, \mathbf{q}_i) \oplus (\mathbf{u}_i, \mathbf{v}_i) \\ &= (\mathbf{q}_i \mathbf{u}_i \mathbf{q}_i^{-1} + \mathbf{p}_i, \mathbf{q}_i \exp(\mathbf{v}_i)).\end{aligned} \tag{2}$$

Here, $\exp(\mathbf{v})$ denotes a 3-dimensional rotation about the axis $\frac{\mathbf{v}}{\|\mathbf{v}\|} \in \mathbb{R}^3$ by angle $\frac{1}{2}\|\mathbf{v}\| \in \mathbb{R}$.

3.2. Multiresolution Representation

The multiresolution representation for a motion signal $\mathbf{M}$ is defined by a series of successively refined signals $\mathbf{M}^{(0)}, \mathbf{M}^{(1)}, \cdots, \mathbf{M}^{(N-1)}$ in addition to a series of displacement maps $\mathbf{D}^{(0)}, \mathbf{D}^{(1)}, \cdots, \mathbf{D}^{(N-1)}$. The construction algorithm starts with the original motion $\mathbf{M} = \mathbf{M}^{(N)}$ to compute a coarser and smoother signal $\mathbf{M}^{(N-1)}$ by reduction, that is, applying a smoothing filter and then removing every other frames to down-sample the signal. One common way to blur signals is through a diffusion process that leads to a local update rule $\mathbf{p}_i \leftarrow \mathbf{p}_i - \lambda L^j \mathbf{p}_i$, where λ is a diffusion coefficient and L is a discrete Laplacian operator. For example, adopting the second Laplacian operator L^2, we have a spatial mask $(-\frac{1}{16}, \frac{4}{16}, \frac{10}{16}, \frac{4}{16}, -\frac{1}{16})$ for smoothing $\mathbf{p}_i$'s, and this mask can also be used for smoothing an orientation signal $\mathbf{q}_i$'s by employing a scheme of Lee [28]. Given a mask $(a_{-k}, \ldots, a_0, \ldots, a_k)$, we smooth a sequence of motion frames, $\{(\mathbf{p}_i, \mathbf{q}_i)\}$ as follows:

$$\mathbf{p}'_i = a_{-k}\mathbf{p}_{i-k} + \cdots + a_0 \mathbf{p}_i + \cdots + a_k \mathbf{p}_{i+k} \tag{3}$$

$$\mathbf{q}'_i = \mathbf{q}_i \exp\left(\sum_{m=-k}^{k-1} b_m \log(\mathbf{q}_{i+m}^{-1} \mathbf{q}_{i+m+1})\right), \tag{4}$$

where

$$b_m = \begin{cases} \sum_{j=m+1}^{k} a_j, & \text{if } 0 \le m \le k-1, \\ \sum_{j=-k}^{m} -a_j, & \text{if } -k \le m < 0. \end{cases}$$

The expansion of $\mathbf{M}^{(N-1)}$ interpolates the missing information through interpolatory subdivision. We use a four-point interpolatory scheme [12, 13] that maps a sequence of motion frames, $\mathbf{M}^{(n-1)} = \{(\mathbf{p}_i^{n-1}, \mathbf{q}_i^{n-1})\}$, to a refined sequence, $\mathbf{M}^{(n)} = \{(\mathbf{p}_i^n, \mathbf{q}_i^n)\}$, where the even numbered frames $(\mathbf{p}_{2i}^n, \mathbf{q}_{2i}^n)$ at level n are the frames $(\mathbf{p}_i^{n-1}, \mathbf{q}_i^{n-1})$ at level $n-1$, and the odd numbered frames are newly inserted between old frames. The subdivision scheme is defined with a subdivision mask $(-\frac{1}{16}, \frac{9}{16}, \frac{9}{16}, -\frac{1}{16})$ that can be generalized for orientation data using Equation (4).

Then, the difference between $\mathbf{M}^{(N)}$ and $\mathbf{M}^{(N-1)}$ is expressed as a displacement map $\mathbf{D}^{(N-1)}$ as follows:

$$\mathbf{D}^{(N-1)} = \mathbf{M}^{(N)} \ominus \mathcal{S}\mathbf{M}^{(N-1)}, \tag{5}$$

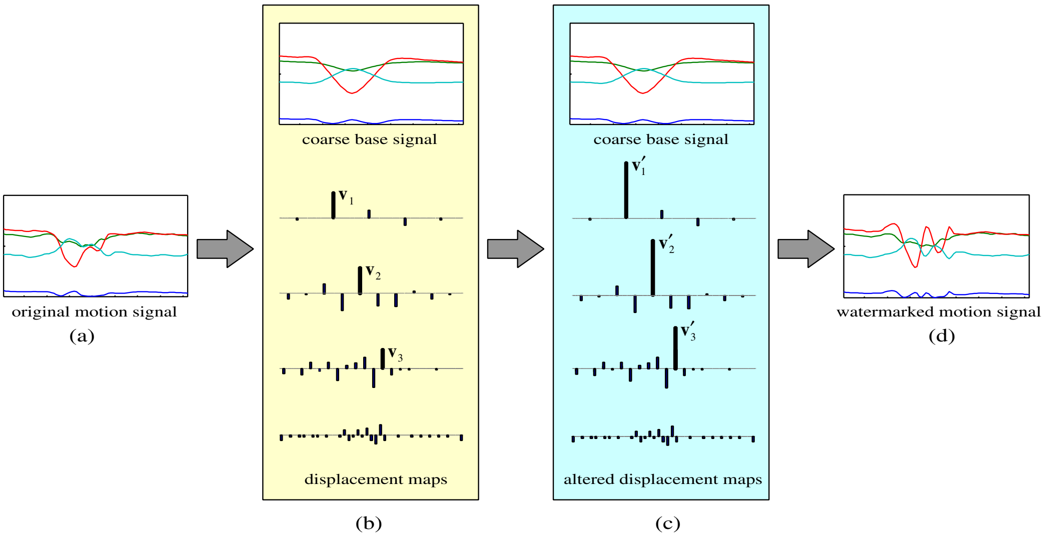

Figure 3: *Overview of our watermarking scheme. Only the orientation component of a motion signal is illustrated. (a) The quaternion signal, $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ of an original signal. (b) The multiresolution representation of the original signal. A coefficient of each displacement map is parameterized by a 3D vector. The displacement maps in the figure show only the first components of those vectors. (c) The multiresolution representation of the watermarked signal. (d) The orientation components of the watermarked signal (see Figure 4) scaled up by a factor of 100.*

where $\mathcal{S}$ is an expansion operator. Cascading these operations until there remains a sufficiently small number of frames in the motion signal, we can construct the multiresolution representation that consists of the coarse base signal $\mathbf{M}^{(0)}$ and a series of displacement maps. The original signal can be reconstructed by cascading expansion and displacement mapping starting from the base signal, that is,

$$\mathbf{M}^{(n)} = \mathcal{S}\mathbf{M}^{(n-1)} \oplus \mathbf{D}^{(n-1)} \qquad (6)$$

for $1 \leq n \leq N$. The original motion must have $(k2^n + 1)$ frames to construct a hierarchy of n levels with its base signal of $(k+1)$ frames. If the original motion has less than $(k2^n + 1)$ frames, then we append extra frames at the end of the signal by duplicating the last frame.

4. Motion Watermarking

For robustness, we adopt a well-known scheme of spread spectrum, that is, embedding a watermark into a wide range of frames in a motion signal. In order to support this scheme, we need a systematic way of identifying the perceptually significant frames scattered over the entire signal. Our multiresolution representation facilitates the scheme due to its capability to manipulate the motion signal at different levels of details. In other words, the multiresolution representation of a motion signal consists of a base signal together with detail coefficients that form a coarse-to-fine hierarchy of motion displacement maps. Each map represents a different detail of the motion. We embed a watermark into the motion by perturbing its detail coefficients. The perturbation of a detail coefficient at a coarse resolution alters a large portion of the motion, while that of a coefficient at a fine resolution affects a small portion. The effect of perturbation reveals a symmetric oscillatory pattern as shown in Figure 4. The size of the pattern depends on the resolution of a displacement map which contains the coefficient to modify. This pattern has a peak value at its center position and rapidly vanishes at both boundaries.

Only the owner of a motion knows the magnitude of alteration for the motion signal due to watermarking, and an attacker may know at best the possible alteration range of magnitude. To completely eliminate a watermark, the attacker is expected to modify the published motion signal as large as possible within this range. Having a different magnitude of a motion signal from the original, the modified one contains noticeable visible defects.

There is a trade-off between the robustness and the imperceptibility of watermarking. A large perturbation is robust but easy to perceive, while a small perturbation is imperceptible but easy to attack. To achieve robustness, it is a common practice to insert a watermark into perceptually significant regions of the signal[8, 34]. Therefore, we embed a watermark into a motion signal by perturbing its large coefficients. To make the watermark imperceptible, we perturb each coefficient with sufficiently small magnitude.

The overview of watermarking is illustrated in Figure 3. In particular, this figure shows how a motion signal is per-

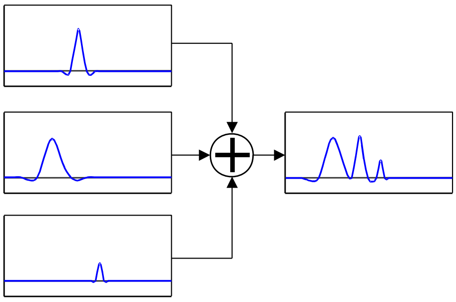

Figure 4: *The perturbation signal: the first components of perturbation applied to three coefficients of the displacement maps.*

turbed for watermarking. An original signal is decomposed into a base motion signal and displacement maps. We select m largest coefficients $\{(\mathbf{u}_1,\mathbf{v}_1), (\mathbf{u}_2,\mathbf{v}_2), \ldots, (\mathbf{u}_m,\mathbf{v}_m)\}$ from the displacement maps. They are altered to $\{(\mathbf{u}'_1,\mathbf{v}'_1), (\mathbf{u}'_2,\mathbf{v}'_2), \ldots, (\mathbf{u}'_m,\mathbf{v}'_m)\}$ by watermarking. Reconstructing the motion signal using these altered displacement maps makes the watermarked signal. The difference between the signal in Figure 3-(a) and that in Figure 3-(d) is shown in figure 4 as a sequence of vectors, $\log(\mathbf{q}_1^{-1}\mathbf{q}'_1), \ldots, \log(\mathbf{q}_n^{-1}\mathbf{q}'_n)$. The magnitude of perturbation is exaggerated for effective explanation.

4.1. Watermarking

Our watermarking scheme consists of following steps: watermark insertion into the original signal, and watermark extraction from the suspected signal, and similarity check between inserted and extracted watermarks.

Watermark generation. Prior to inserting watermark into a motion signal, we first generate a watermark $\mathbf{w} = \{w_1, \ldots, w_m\}$, where $w_i \in \mathbb{R}$. Each w_i is sampled independently from a normal distribution with zero mean and unit variance. To make the watermarking scheme be non-invertible, we apply a cryptographic hash function, MD5[38] to the original motion signal concatenated with the owner's key, and then seed a random number generator with the hashed value[9]. Given the m largest coefficients for watermarking, each element of the watermark is used to perturb one of those coefficients.

Watermark insertion and extraction. We insert $\mathbf{w}$ into the original signal $\mathbf{M}$ in order to obtain a watermarked one $\mathbf{M}'$. After selecting the m largest coefficients, we scale each of the coefficients using $\mathbf{w}$. Let $\mathbf{D}^{(l)}$ be the displacement map at level l of the multiresolution representation of $\mathbf{M}$, and $(\mathbf{u}_i,\mathbf{v}_i)$ the i-th largest coefficient of the displacement maps, $\{\mathbf{D}^{(0)},\mathbf{D}^{(1)},\ldots,\mathbf{D}^{(N-1)}\}$, where $N-1$ is the maximum level of hierarchy for the displacement maps. We compute the perturbed displacement vector $(\mathbf{u}'_i,\mathbf{v}'_i)$ for $i = 1,\ldots,m$ as follows:

$$(\mathbf{u}'_i,\mathbf{v}'_i) = (1+\alpha w_i)(\mathbf{u}_i,\mathbf{v}_i), \tag{7}$$

where α is a user-provided scaling parameter. This perturbation gives the new displacement maps $\{\mathbf{D}'^{(0)},\mathbf{D}'^{(1)},\ldots,\mathbf{D}'^{(N-1)}\}$. We construct the watermarked signal $\mathbf{M}'$ from these displacement maps by successively adding them to the base signal. As long as the selected displacement vector is not a zero vector, the insertion process is invertible.

Watermark extraction is the inverse of watermark insertion in some sense. From the suspected motion $\mathbf{M}^*$, we build its multiresolution representation and extract the watermark $\mathbf{w}^*$ by computing w that minimize the following function $f(w)$:

$$f(w) = \|(\mathbf{u}^*_i,\mathbf{v}^*_i) - (1+\alpha w)(\mathbf{u}_i,\mathbf{v}_i)\|^2, \tag{8}$$

where α, $(\mathbf{u}_i,\mathbf{v}_i)$, $(\mathbf{u}^*_i,\mathbf{v}^*_i)$ is the same as defined in Equation (7). This function is minimized at

$$w = \frac{(\mathbf{u}^*_i,\mathbf{v}^*_i)\cdot(\mathbf{u}_i,\mathbf{v}_i) - \|(\mathbf{u}_i,\mathbf{v}_i)\|^2}{\alpha\|(\mathbf{u}_i,\mathbf{v}_i)\|^2}, \tag{9}$$

which becomes the extracted watermark w^*_i.

Analysis of similarity. A watermarking scheme must minimize the false-positive probability, that is, the probability of incorrectly asserting that the data is watermarked. The false-negative probability is the probability of failing to detect watermarked data. Ideally, it is desired to minimize both false-positive and false-negative probabilities, simultaneously. However, we have a trade-off between them; enforcing low false-negative probability often leads to high false-positive probability to yield the weak fidelity of an ownership claim. Therefore, like other well-known watermarking schemes[7, 8, 34], we focus on decreasing the false-positive probability.

We compare the inserted and extracted watermarks statistically. At this point, we remove outliers, that is, elements of which differences from the mean value are greater than a given threshold. Then, as in Praun *et al.*[34], we compute linear correlation between watermarks:

$$\rho = \frac{\sum_i (w^*_i - \overline{w^*})(w_i - \overline{w})}{\sqrt{\sum_i (w^*_i - \overline{w^*})^2 \times \sum_i (w_i - \overline{w})^2}}, \tag{10}$$

where w_i and w^*_i are an element of the original watermark and that of the extracted watermark after discarding outliers, respectively. In addition, $\overline{w}$ and $\overline{w^*}$ are the average of the elements in $\mathbf{w}$ and that of the elements in $\mathbf{w}^*$, respectively. Finally, we compute the *false-positive* probability using Student's t-test[35].

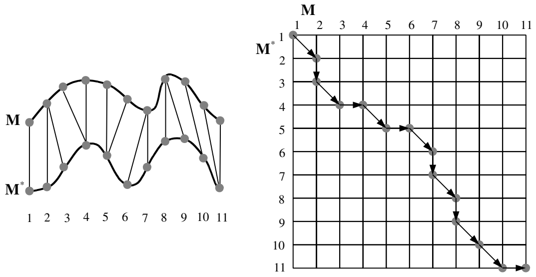

Figure 5: *Representation of signal resampling.*

4.2. Motion Signal Aligning

A suspected signal may be cropped or non-uniformly scaled along the time axis. Therefore, before inspecting the watermark, we optionally align the suspected signal to the original. Let $\mathbf{M}^*$ and $\mathbf{M}$ be the suspected and original signals, respectively. That is,

$$\begin{aligned} \mathbf{M}^* &= \{(\mathbf{p}_1^*,\mathbf{q}_1^*),(\mathbf{p}_2^*,\mathbf{q}_2^*),...,(\mathbf{p}_l^*,\mathbf{q}_l^*)\}, \\ \mathbf{M} &= \{(\mathbf{p}_1,\mathbf{q}_1),(\mathbf{p}_2,\mathbf{q}_2),...,(\mathbf{p}_n,\mathbf{q}_n)\}. \end{aligned} \tag{11}$$

Resampling. Suppose that the suspected signal $\mathbf{M}^*$ is non-uniformly scaled along the time axis. Then, we resample $\mathbf{M}^*$ to align with $\mathbf{M}$ using the dynamic time warping scheme[5, 22, 41]. Let $\mathbf{G}$ be a $l \times n$ rectangular grid, where the column indices correspond to the frame numbers of $\mathbf{M}$ and the row indices do to those of $\mathbf{M}^*$ (see Figure 5). Thus, a grid point (i, j) represents an ordered pair of frames, that is, the i-th frame of $\mathbf{M}^*$ and the j-th frame of $\mathbf{M}$. With this grid, the resampling problem is reduced to constructing a path on the grid $\mathbf{G}$ from $(1,1)$ to (l,n). At each grid point, the path is allowed to move horizontally, vertically, or diagonally. To minimize the total difference between $\mathbf{M}$ and $\mathbf{M}^*$, a dynamic programming technique is employed. Letting $M_{i,j}$ be the minimum difference between the subsignal of $\mathbf{M}^*$ up to the i-th frame and that of $\mathbf{M}$ up to the j-th frame, we have

$$M_{i,j} = \min(M_{i-1,j-1}, M_{i-1,j}, M_{i,j-1}) + d_{i,j}, \tag{12}$$

where $d_{i,j} = \|(\mathbf{p}_i^*,\mathbf{q}_i^*) \ominus (\mathbf{p}_j,\mathbf{q}_j)\|$, that is, the difference between the i-th frame of $\mathbf{M}^*$ and the j-th frame of $\mathbf{M}$. In addition, $M_{1,1} = d_{1,1}$ and $M_{i,j} = \infty$ for $i < 1$, $j < 1$, $i > l$, or $j > n$. Clearly, $M_{l,n}$ can be obtained in $O(ln)$ time, and thus we can optimally resample $\mathbf{M}^*$ in the same time bound.

Registration. Now, suppose that $\mathbf{M}^*$ is a cropped version of $\mathbf{M}$. Then, the frames in the cropped portion of $\mathbf{M}$ does not correspond to any frame of $\mathbf{M}^*$. Therefore, we resample $\mathbf{M}^*$ to register with $\mathbf{M}$ and fill the missing portions of $\mathbf{M}^*$ using their corresponding portions of $\mathbf{M}$. We examine every portion of $\mathbf{M}$, that is, the subsequence of consecutive frames of $\mathbf{M}$ to find an optimal portion that minimizes its difference from $\mathbf{M}^*$, using the dynamic time warping. Since there are $\frac{n(n-1)}{2}$ possible portions of $\mathbf{M}$, and the difference between each of them and $\mathbf{M}^*$ can be computed in $O(nl)$ time, we can register $\mathbf{M}^*$ with $\mathbf{M}$ in $O(n^3 l)$ time.

Table 1: *Watermark detection results for four motion data, DataA: Fly Spin Kick, DataB: Back Flip, DataC: Broad Jump, and DataD: Blown Back*

	Attack	DataA	DataB	DataC	DataD
1.	No attack	$\mathbf{10^{-99}}$	$\mathbf{10^{-99}}$	$\mathbf{10^{-99}}$	$\mathbf{10^{-99}}$
2.	Noise 0.2%	$\mathbf{10^{-18}}$	10^{-17}	10^{-52}	10^{-26}
3.	Noise 0.7%	$\mathbf{10^{-9}}$	10^{-6}	10^{-21}	10^{-16}
4.	Cropping	10^{-35}	$\mathbf{10^{-22}}$	10^{-14}	10^{-9}
5.	Smoothing 1	10^{-24}	10^{-9}	10^{-15}	10^{-11}
6.	Smoothing 2	10^{-7}	10^{-5}	10^{-5}	$\mathbf{10^{-12}}$
7.	Simplifying 1	10^{-35}	10^{-35}	$\mathbf{10^{-35}}$	10^{-35}
8.	Simplifying 2	10^{-27}	10^{-38}	$\mathbf{10^{-27}}$	10^{-17}
9.	Simplifying 3	10^{-15}	10^{-28}	$\mathbf{10^{-32}}$	10^{-15}
10.	Time warping 1	10^{-8}	10^{-27}	10^{-13}	$\mathbf{10^{-24}}$
11.	Time warping 2	10^{-10}	10^{-3}	10^{-30}	$\mathbf{10^{-36}}$
12.	Enhancement	10^{-4}	$\mathbf{10^{-4}}$	10^{-12}	10^{-4}
13.	Attenuation	10^{-7}	$\mathbf{10^{-5}}$	10^{-14}	10^{-15}
14.	2nd watermark	$\mathbf{10^{-73}}$	10^{-80}	10^{-74}	10^{-76}

5. Experimental Results

In this section, we conduct some experiments to show the effectiveness of our watermarking scheme. We use four motion clips, *Fly Spin Kick*, *Back Flip*, *Broad Jump*, and *Blown Back* as shown in Figure 6. Since watermarking depends on random numbers, we perform experiments five times for each motion clip using different watermarks and present the median of false-positive probabilities as their estimation. We use 10^{-4} as the scaling parameter α for pelvis and 10^{-2} for the others. Those scaling parameters are determined experimentally to provide robustness while still keeping imperceptibility. The length of a watermark for a motion signal can be adjusted to the characteristics of the signal. In our experiments, the length is chosen to be 20 for all motion clips. To extract the watermark from a distorted signal, we use 2.5 times of its variance as the threshold to remove outliers in the extracted watermark.

Table 1 summarizes our experimental results for a variety of attacks on four motion clips. Each entry shows the estimation of the false-positive probability for an attack on a motion signal corresponds to the head of trajectory in a motion clip. A row has four entries showing experimental results for the attack given in the leftmost entry. Figure 7 exhibits the motion data corresponding to selected table entries written in bold face. The figures in the first row show the original motion data.

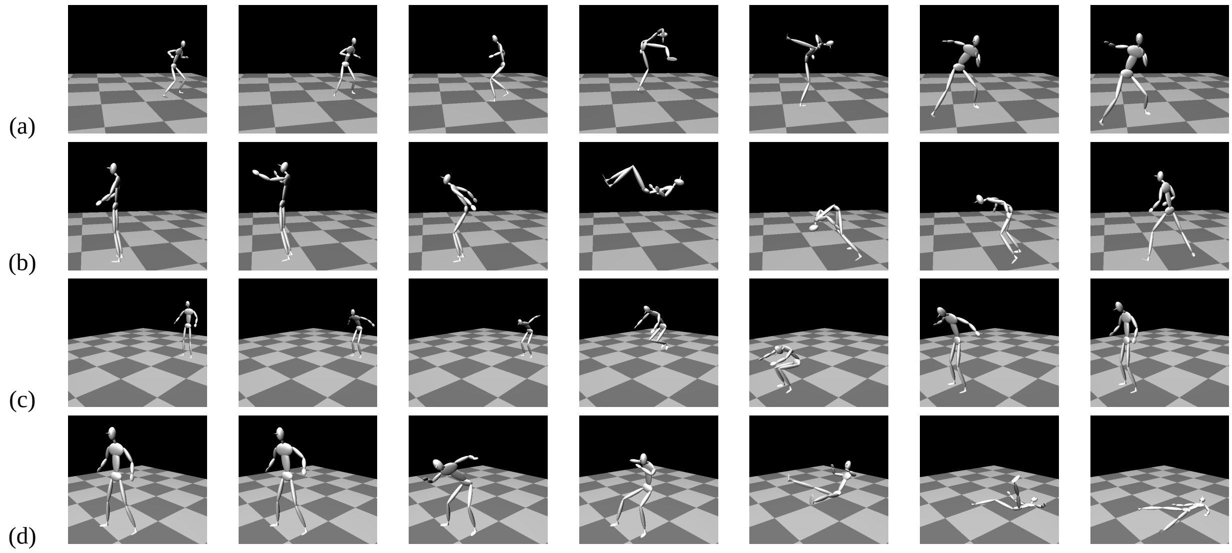

Figure 6: *Four original unwatermarked motion data. (a) Fly Spin Kick. (b) Back Flip. (c) Broad Jump. (d) Blown Back.*

No attack. The first row shows the results for watermarked signals without any attack. Their estimated false-positive probabilities are almost but not exactly zero due to precision errors. The distortion due to precision errors is similar to that caused by a small perturbation.

Noise. Rows 2 and 3 address the cases of 0.2% and 0.7% white noise attacks, respectively. Here, each percentage represents the ratio between the largest amplitude of noise and the largest rotation angle in the motion signal. The noise is added to the signal.

Cropping. Row 4 demonstrates the ability of our scheme for cropped signal. Though the portion of a watermark in a cropped portion is eliminated completely, its remaining portion can be detected with little obstruction.

Smoothing. Rows 5 and 6 present the results after applying a smoothing filter three times. We use the averaging filter of unit radius for row 5:

$$\mathcal{H}^A(\mathbf{q}_i) = \frac{\mathbf{q}_{i-1} + \mathbf{q}_i + \mathbf{q}_{i+1}}{\|\mathbf{q}_{i-1} + \mathbf{q}_i + \mathbf{q}_{i+1}\|}. \tag{13}$$

For row 6, we use the spatial filter of Lee [28]:

$$\mathcal{H}^S(\mathbf{q}_i) = \mathbf{q}_i \exp\left(\frac{\lambda}{24}(\omega_{i-2} - 3\omega_{i-1} + 3\omega_i - \omega_{i+1})\right), \tag{14}$$

where ω_i is a rotation vector from the i-th frame to the $(i+1)$-th frame and λ is a damping factor that controls the rate of convergence. We set a damping factor λ to 1.0. Since the smoothing filters change fine details of motion signals, our scheme is resilient to this attack.

Simplifying. We test our watermarking scheme for a simplified signal whose fine levels of their multiresolution representation are eliminated. Rows 7, 8, and 9 show the results for a signal with level 0 eliminated, levels 0 and 1 eliminated, and levels 0, 1 and 2 eliminated, respectively. This attack may destroy parts of a watermark. However, our scheme can detect the remaining parts of the watermark embedded into the coarser levels.

Time warping. Rows 10 and 11 address an attack due to time warping. The attack of uniform scaling with a factor of 0.5 is presented in row 10, and that of non-uniform scaling is shown in row 11. Since the watermark is spread widely in the motion signal, our scheme is robust for this type of attack.

Enhancement and Attenuation. Rows 12 and 13 show the results for the enhanced and attenuated versions, respectively. For row 12, we multiply a constant factor of 1.5 to each of the detail coefficients at the coarsest level and its next finer level to enhance motion data. For row 13, we use a constant factor of 0.6 for attenuation. Each of these attacks alters parts of the embedded watermark as in the simplifying attack. Therefore, we can similarly detect the remaining watermark.

2nd watermarking. The resilience under a second watermarking is shown in row 14. We insert a watermark to the watermarked signal. This watermark is different from that for the original signal. Such an attack resembles that of adding noise to the displacement vectors of the signal.

For each of experiments, its false-positive probability is sufficiently low to support the robustness of our scheme.

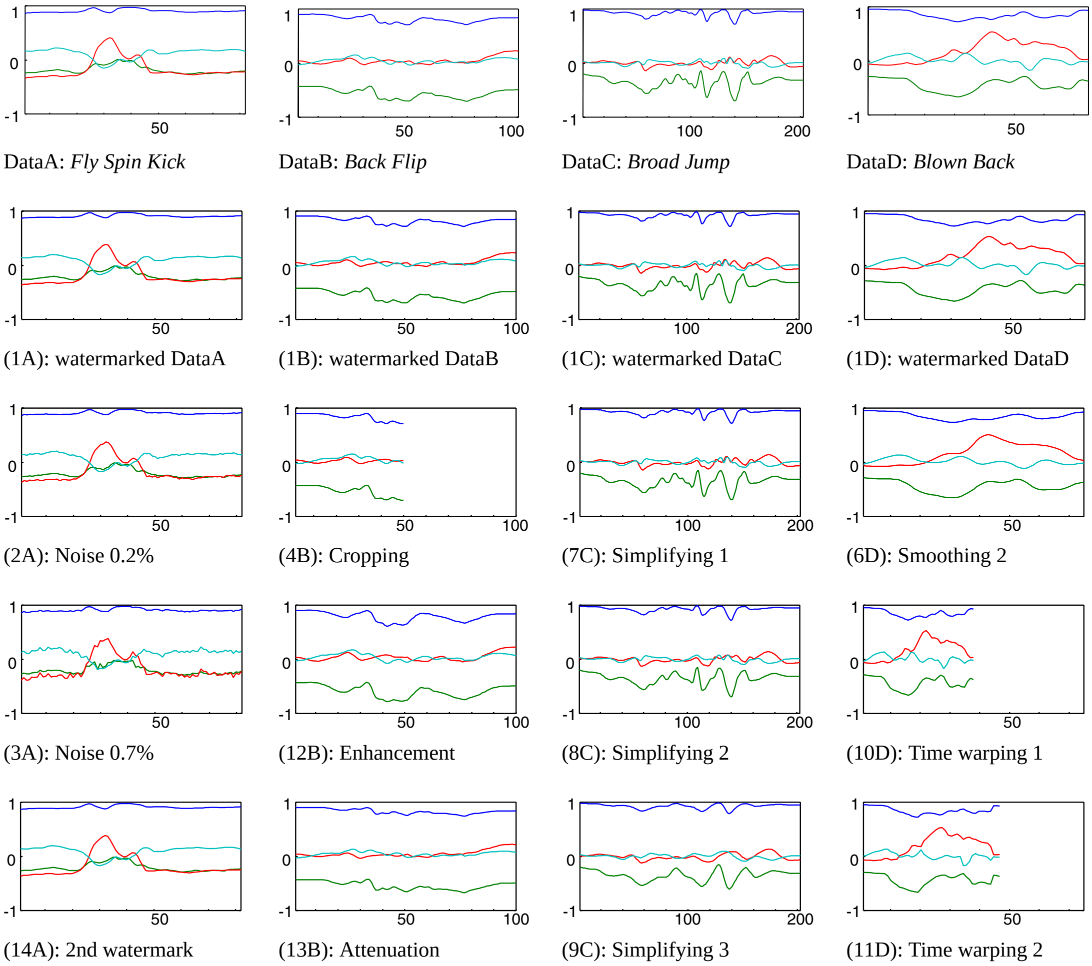

Figure 7: *Original and watermarked motion signals (top two rows), and various attacks*

6. Conclusion

In this paper, we have presented a practical method for embedding a watermark into a motion signal. The key idea of our scheme is to combine two novel ideas: multiresolution representation and spread spectrum. For robustness, we adopt the idea of spread spectrum, that is, embedding a watermark into a wide range of frames. This idea is instantiated through the multiresolution representation of a motion signal, by identifying the perceptually significant frames of the signal to embed a watermark. Experimental results show that our scheme is robust to various signal processing operations or motion editing such as enhancement, attenuation, and time warping. In the future, we would like to improve our motion watermarking scheme to be resilient to more complex motion editing[27] such as non-uniform scaling along the amplitude axis.

7. Acknowledgements

This work was supported by NRL (National Research Laboratory) project of KISTEP (Korea Institute of Science & Technology Evaluation and Planning).

References

1. W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM Systems Journal*, 35(3&4):313–336, 1996.
2. D. Berman, J. Bartell, and D. Salesin. Multiresolution painting and compositing. *Computer Graphics (Proceedings of SIGGRAPH 94)*, pages 85–90, July 1994.
3. J. S. De Bonet. Multiresolution sampling procedure for analysis and synthesis of texture images. *Computer Graphics (Proceedings of SIGGRAPH 97)*, pages 361–368, August 1997.

4. L. Boney, A. H. Tewfik, and K. N. Hamdy. Digital watermarks for audio signals. In *Proceedings of the Third IEEE International Conference on Multimedia Computing Systems, 1996*, pages 473–480, 1996.

5. A. Bruderlin and L. Williams. Motion signal processing. In *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 97–104, August 1995.

6. A. Certain, J. Popović, T. DeRose, T. Duchamp, D. Salesin, and W. Stuetzle. Interactive multiresolution surface viewing. *Computer Graphics (Proceedings of SIGGRAPH 96)*, pages 97–115, August 1996.

7. I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio, and video. In *IEEE Proceedings of International Conference on Image Processing*, volume 3, pages 243–246, 1996.

8. I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing*, 6(12):1673–1687, December 1997.

9. S. Craver, N. Memon, and B. Yeo. Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications. *IEEE Journal on Selected Areas in Communications*, 16(4):573–586, May 1998.

10. R. A. DeVore, B. Jawerth, and B. J. Lucier. Image compression through wavelet transform coding. *IEEE Transactions on Information Theory*, 38(2):719–746, 1992.

11. R. A. DeVore, B. Jawerth, and B. J. Lucier. Surface compression. *Computer Aided Geometric Design*, 9(3):219–239, 1992.

12. S. Dubuc. Interpolation through an iterative scheme. *J. Math. Anal. Appl.*, 114:185–204, 1986.

13. N. Dyn, J. Gregory, and D. Levin. A 4-point interpolatory subdivision scheme for curve design. *Computer Aided Geometric Design*, 7:129–140, 1987.

14. M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, and W. Stuetzle. Multiresolution analysis of arbitrary meshes. *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 173–182, August 1995.

15. A. Finkelstein, C. E. Jacobs, and D. H. Salesin. Multiresolution video. *Computer Graphics (Proceedings of SIGGRAPH 96)*, pages 281–290, August 1996.

16. A. Finkelstein and D. H. Salesin. Multiresolution curves. *Computer Graphics (Proceedings of SIGGRAPH 94)*, pages 261–268, July 1994.

17. S. J. Gortler and M. F. Cohen. Hierarchical and variational geometric modeling with wavelets. In *Proceedings of the 1995 ACM Symposium on Interactive 3D Graphics*, pages 35–42, 1995.

18. S. J. Gortler, P. Schröder, M. F. Cohen, and P. Hanrahan. Wavelet radiosity. *Computer Graphics (Proceedings of SIGGRAPH 93)*, pages 221–230, August 1993.

19. I. Guskov, W. Sweldens, and P. Schröder. Multiresolution signal processing for meshes. *Computer Graphics (Proceedings of SIGGRAPH 99)*, pages 325–334, August 1999.

20. F. Hartung and B. Girod. Fast public-key watermarking of compressed video. In *Proceedings of International Conference on Image Processing*, volume 1, pages 528–531, 1997.

21. D. J. Heeger and J. R. Bergen. Pyramid based texture analysis/synthesis. *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 229–238, August 1995.

22. X. D. Huang, Y. Ariki, and M. A. Jack. *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, 1990.

23. C. E. Jacobs, A. Finkelstein, and D. H. Salesin. Fast multiresolution image querying. *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 277–286, August 1995.

24. L. Kobbelt, S. Campagna, J. Vorsatz, and H.-P. Seidel. Interactive multi-resolution modeling on arbitrary meshes. *Computer Graphics (Proceedings of SIGGRAPH 98)*, pages 105–114, July 1998.

25. L. Kobbelt and P. Schröder. A multiresolution framework for variational subdivision. *ACM Transactions on Graphics*, 17(4):209–237, 1998.

26. A. W. F. Lee, W. Sweldens, P. Schröder, L. Cowsar, and D. Dobkin. MAPS: Multiresolution adaptive parameterization of surfaces. *Computer Graphics (Proceedings of SIGGRAPH 98)*, pages 95–104, July 1998.

27. J. Lee and S. Y. Shin. A hierarchical approach to interactive motion editing for human-like figures. In *Computer Graphics (Proceedings of SIGGRAPH 99)*, pages 39–48, August 1999.

28. Jehee Lee. *A Hierarchical Approach to Motion Analysis and Synthesis for Articulated Figures*. PhD thesis, *DCS955295*, Department of Computer Science, KAIST, 2000.

29. Z. Liu, S. G. Gortler, , and M. F. Cohen. Hierarchical spacetime control. In *Computer Graphics (Proceedings of SIGGRAPH 94)*, pages 35–42, July 1994.

30. M. Lounsbery, T. D. DeRose, and J. Warren. Multiresolution analysis for surfaces of arbitrary topological type. *ACM Transactions on Graphics*, 16(1):34–73, 1997.

31. R. L. Peterson, R. E. Ziemer, and D. E. Borth. *Introduction to Spread Spectrum Communications*. Prentice Hall, Inc., A Simon & Schuster Company Englewood Cliffs, New Jersey 07632, 1995.

32. R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread-spectrum communications – a tutorial. *IEEE Transactions on Communications*, COM-30(5):855–884, May 1982.

33. C. I. Podilchuk and W. Zeng. Image-adaptive watermarking using visual models. *IEEE Journal on Selected Areas in Communications*, 16(4):525–539, May 1998.

34. E. Praun, H. Hoppe, and A. Finkelstein. Robust mesh watermarking. In *Computer Graphics (Proceedings of SIGGRAPH 99)*, pages 49–56, August 1999.

35. W. H. Press, S. A. Teulkolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, 1996.

36. J. J. K. Ó. Ruanaidh, F. M. Boland, and O. Sinnen. Watermarking digital images for copyright protection. In *Proceedings of the Electronic Imaging and Visual Arts Conference*, pages 243–246, February 1996.

37. J. J. K. Ó. Ruanaidh, W. J. Dowling, and F. M. Boland. Phase watermarking of digital images. In *IEEE Proceedings of International Conference on Image Processing*, pages 239–242, September 1996.

38. B. Schneier. *Applied Cryptography: protocols, algorithms, and source code in C*. John Wiley & Sons, Inc., 2nd edition, 1996.

39. P. Schröder and W. Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 161–172, August 1995.

40. R. G. V. Schyndel, A. Z. Tirkel, and C. F. Osborne. A digital watermark. In *IEEE Proceedings of International Conference on Image Processing*, volume 2, pages 86–90, 1994.

41. T. W. Sederberg and E. Greenwood. A physically based approach to 2-D shape blending. In *Proceedings of the 19th annual conference on Computer graphics*, pages 25–34, July 27–31 1992.

42. K. Tanaka, Y. Nakamura, and K. Matsui. Embedding secret information into a dithered multi-level image. In *MILCOM'90: A New Era. 1990 IEEE Military Communications Conference*, volume 1, pages 216–220, 1990.

43. G. Voyatzis and I. Pitas. The user of watermarks in the protection of digital multimedia products. In *Proceedings of IEEE*, volume 87, pages 1197–1207, 1999.

44. A. Witkin and Z. Popović. Motion warping. In *Computer Graphics (Proceedings of SIGGRAPH 95)*, pages 105–108, August 1995.

45. D. Zorin, P. Schröder, and W. Sweldens. Interactive multiresolution mesh editing. *Computer Graphics (Proceedings of SIGGRAPH 97)*, pages 259–268, August 1997.