AFFINE INVARIANT WAVELET TRANSFORM

Victor H. S. Ha and José M. F. Moura

Carnegie Mellon University
Department of Electrical and Computer Engineering
5000 Forbes Avenue
Pittsburgh, PA 15213-3890
{hvh, moura}@ece.cmu.edu

ABSTRACT

We present a two-dimensional wavelet transform that is invariant to affine distortions of the input signal. Affine distortions include geometric effects such as translation, reflection, uniform and anisotropic scaling, rotation, and shearing of the input signal. Invariance of the wavelet transform to affine distortions is achieved in our work by developing an algorithm that reduces replicas of a signal related by affine distortions to a unique prototype signal. The affine invariant wavelet transform is then defined as the two-dimensional wavelet transform of the prototype signal, which provides the wavelet coefficients that are invariant to affine distortions of the input signal. We describe our algorithm and show examples that demonstrate our claims.

1. INTRODUCTION

The wavelet transform is widely used in many interesting signal processing applications including image coding and analysis. The discrete wavelet transform in particular is very popular due to its compact representation and efficient implementation. The one dimensional discrete wavelet transform of a square integrable signal $f(t) \in L^2(\mathbb{R})$ is written as

$$DWT\{f(t)\} = \langle f(t), \psi_{j,n}(t)\rangle = \int_{-\infty}^{\infty} f(t)\psi^*_{j,n}(t)dt \quad (1)$$

where the wavelet frame $\{\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}}\psi(\frac{t-2^j n}{2^j})\}_{(j,n)\in\mathbb{Z}^2}$ forms an orthonormal basis for $L^2(\mathbb{R})$.

Unfortunately, the discrete wavelet transform lacks shiftability as discussed in [1]. That is, the wavelet coefficients vary significantly when the input signal is shifted, for example, by translation, scale, and rotation. This lack of shiftability is troublesome especially in image identification and classification problems in which objects present in the image are affected by geometric distortions resulting from translation, scaling, rotation, and shearing. Research has been reported on the shiftability of the wavelet transform. Translation invariant wavelet and wavelet packet transforms are studied in [2] [3] [4] [5]. Joint shiftability in translation and scale is considered in [1] by relaxing the shiftability constraint in one of the domains, and in [6] by iterating back and forth between shiftability invariance on the two domains. A non-linear approach is investigated in [7] for joint shiftability under the *similarity* transform, which is restricted to translation, uniform scaling, and rotation.

In this paper, we present a wavelet transform that is invariant to affine distortions. Affine distortions arise in many imaging environments where the objects in the scene are in relative motion with respect to the imaging sensor. Affine distortions include translation, reflection, uniform and anisotropic scaling, rotation and shearing. Anisotropic scaling and shearing, in particular, change the shape of the signal and pose a significant challenge in achieving affine invariance. Our affine invariant wavelet transform (AIWT) generates unique transform coefficients for arbitrary affine distorted versions of the same signal.

In the next section, we introduce the concept of the affine invariant wavelet transform. In section 3, we describe the blind normalization algorithm (BNA) that provides the affine invariance. Experimental results are shown in section 4. Section 5 concludes the paper. Proofs are ommitted due to lack of space. They will be presented in the full version of the paper.

2. AFFINE INVARIANT WAVELET TRANSFORM

A two-dimensional (2D) signal $f(x, y)$ is distorted by an affine transform into the signal $f^d(x, y)$ where

$$f^d(x, y) = f(\bar{x}, \bar{y}). \quad (2)$$

and

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} A_1^1 & A_2^1 \\ A_1^2 & A_2^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \mathbf{A}\mathbf{x} + \mathbf{t} \quad (3)$$

for $\{A_i^j\}_{(i,j)\in\{1,2\}} \in \mathbb{R}$ and $\{x_t, y_t\} \in \mathbb{R}$.

By recentering the signal with respect to the center of mass, $\{g_x, g_{\bar{x}}\}$, we remove the translation factor $\mathbf{t} = [x_t, y_t]^T$ as in

$$\bar{\mathbf{x}}^* = (\bar{\mathbf{x}} - \mathbf{g}_{\bar{\mathbf{x}}}) = \mathbf{A}(\mathbf{x} - \mathbf{g}_{\mathbf{x}}) = \mathbf{A}\mathbf{x}^* \quad (4)$$

and safely disregard the translation distortion of the signal $f^d(x, y)$. Therefore, we consider only the linear part of the affine transform in the remainder of the paper.

Our goal with the affine invariant wavelet transform is to generate the same set of transform coefficients $\mathcal{C}$ for any affine distorted signal $f^d(x, y)$. Denote by $AIWT\{\cdot\}$ the affine invariant wavelet transform. Then, for a 2D signal $f(x, y)$ and its affine distorted version $f^d(x, y)$,

$$\mathcal{C} = \text{AIWT}\{f(x, y)\} = \text{AIWT}\{f^d(x, y)\} = \mathcal{C}^d. \quad (5)$$

The AIWT allows us to represent all affine distorted signals by the same set of wavelet coefficients. To describe the basic idea underlying the AIWT, we define $\mathcal{D}(f)$ to be the set of all affine distorted versions of the signal f, and let f^o be a representative prototype of the set $\mathcal{D}(f)$. Given an affine distorted signal $f^d \in \mathcal{D}$, we reduce the signal f^d to the prototype signal f^o by the blind normalization algorithm (BNA) that we describe in section 3. The normalization algorithm is blind because it does not know either the prototype signal f^o of the set $\mathcal{D}$ to which f^d belongs, or the particular affine transform $\mathbf{A}$ that relates f^d and f^o. Then, the affine invariant wavelet transform of f^d becomes simply the wavelet transform of the prototype signal f^o, i.e.,

$$\text{AIWT}\{f^d\} = \text{WT}\{f^o\}. \tag{6}$$

Due to f^o being unique, this results in a unique set of wavelet coefficients for the set $\mathcal{D}(f)$.

In the next section, we introduce the blind normalization algorithm that converts f^d to f^o.

3. THE BLIND NORMALIZATION ALGORITHM (BNA)

The blind normalization algorithm (BNA) transforms an affine distorted signal f^d into a prototype signal f^o. The algorithm does not compute the affine parameters $\{A_j^i\}$, thus no explicit inversion of the affine matrix is performed. Further, the BNA uses no knowledge of the prototype signal f^o.

The BNA consists of two components. The first is a Rotate and Scale (RnS) step that rotates the signal by a fixed angle θ followed by a scale normalization. The second component is the computation of the orientation indicator index (OII) that is defined below. The algorithm iterates repeatedly the RnS step and computes the OII after each iteration. After a finite number of iterations, the signal corresponding to the maximum value of the OII is chosen and stored as the normalized signal f^n. Since the BNA reduces any affine distorted replica f^d to the prototype f^o, the normalized signal f^n is identical to the prototype signal f^o. We then compute the wavelet transform of the normalized signal f^n to generate the AIWT coefficients. We now detail the two components of the BNA: RnS and OII.

3.1. Rotate and Scale (RnS)

The Rotate and Scale (RnS) step restores the standard shape of the prototype signal from any arbitrary affine distorted signal. In this step, the signal is rotated by a fixed angle θ and is scaled by factors $\{1/\alpha, 1/\beta\}$, where α and β are the dimensions of the region of support of the signal along the x– and y– axes defined as

$$\begin{aligned}\alpha &= \Big|\max_x\{\, x : f(x,y) \neq 0 \text{ for some } y \in \mathbb{R}\} \\ &\quad - \min_x\{\, x : f(x,y) \neq 0 \text{ for some } y \in \mathbb{R}\}\Big| \qquad (7)\\ \beta &= \Big|\max_y\{\, y : f(x,y) \neq 0 \text{ for some } x \in \mathbb{R}\} \\ &\quad - \min_y\{\, y : f(x,y) \neq 0 \text{ for some } x \in \mathbb{R}\}\Big|.\end{aligned}$$

If the signal $f(x,y)$ is the input to the RnS step, the outcome $f(\bar{x},\bar{y})$ is obtained by a coordinate transform,

$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha} & 0 \\ 0 & \frac{1}{\beta} \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{8}$$

The above matrix multiplication is repeated N times in the BNA

$$\bar{\mathbf{x}} = (\mathbf{S_N R}) \ldots (\mathbf{S_1 R})(\mathbf{S_o R})\mathbf{x} \tag{9}$$

where the matrix $\mathbf{R}$ is a rotation by a fixed angle θ and the scaling matrix $\mathbf{S_i}$ contains the scaling factors $\{1/\alpha_i, 1/\beta_i\}$ that are recomputed for each rotated signal. We call the multiplication by the scaling matrix $\mathbf{S_i}$ the scale normalization. The number of repetitions N is determined by the choice of the rotation angle θ, which repeats itself after $N = 2m$ iterations if $\theta = \pi/m$.

The outcome of the RnS step is the shape-recovered signal at different orientations. In order to obtain the one at the standard orientation of the prototype signal f^o, we need the orientation indicator index (OII) that monitors the orientation of the signal. We introduce this new index measure next.

3.2. Orientation Indicator Index (OII)

We define the orientation indicator index (OII) for the signal $f(x,y)$

$$\begin{aligned}\text{OII}_f &= \sqrt{\mu_x^2 + \mu_y^2} \qquad (10)\\ \mu_x &= \int_{\mathcal{R}} (x - g_x)^3 f(x,y)\, dx\, dy \qquad (11)\\ \mu_y &= \int_{\mathcal{R}} (y - g_y)^3 f(x,y)\, dx\, dy. \qquad (12)\end{aligned}$$

Because the signal $f(x,y)$ is centered at its center of mass by Equation (4), we rewrite Equation (10)

$$\begin{aligned}[\text{OII}_f]^2 &= \left[\int_{\mathcal{R}} x^3 f(x,y) dx dy\right]^2 + \left[\int_{\mathcal{R}} y^3 f(x,y) dx dy\right]^2 \qquad (13)\\ &= \iint_{\mathcal{R}} (x^3u^3 + y^3v^3) f(x,y) f(u,v) dx dy du dv.\end{aligned}$$

If the signal $f(x,y)$ is rotated by an angle θ, the rotated signal $f^\theta(x,y)$ is equal to $f(\bar{x},\bar{y})$ where $\bar{\mathbf{x}} = \mathbf{R}\mathbf{x}$. The orientation indicator index of f^θ is

$$\begin{aligned}\text{OII}_{f^\theta} &= \sqrt{\mu_{\bar{x}}^2 + \mu_{\bar{y}}^2} \qquad (14)\\ \mu_{\bar{x}} &= \int_{\mathcal{R}} (\cos(\theta)\, x + \sin(\theta)\, y)^3 f(x,y)\, dx\, dy \qquad (15)\\ \mu_{\bar{y}} &= \int_{\mathcal{R}} (-\sin(\theta)\, x + \cos(\theta)\, y)^3 f(x,y)\, dx\, dy. \qquad (16)\end{aligned}$$

From Equations (10) – (16), we derive the following properties.

Properties of OII:

1. If $f(x,y)$ is symmetric about both the x– and y– axes, the OII is zero for all rotation angles $\theta = [0, 2\pi]$. This is because the third order central moments $\{\mu_x, \mu_y\}$ are zero for such a signal.
2. The OII is periodic with period $\pi/2$. That is, $\text{OII}_{f^\theta} = \text{OII}_{f^{(\theta+\frac{\pi}{2})}}$. This is due to the relationship,

$$\begin{bmatrix} \mu_{\bar{x}} \\ \mu_{\bar{y}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \tag{17}$$

for any two signals separated by a rotation of $\pi/2$ radians.

3. If the signal is symmetric about only one or none of the x– and y–axes, the OII has a finite number of maximum points in each interval of $\theta = [0, 2\pi]$. The maximum points in the OII plot represent unique orientations of the signal.

We now present the blind normalization algorithm.

The Blind Normalization Algorithm:

1. Recentering: Recenter $f^d(x, y)$ with respect to the center of mass $\{g_{\bar{x}}, g_{\bar{y}}\}$.
2. Scale Normalization: Compute $\{\alpha, \beta\}$ from Equation (7). Scale the pixel coordinates (x, y) to $(\frac{x}{\alpha}, \frac{y}{\beta})$.
3. Rotate and Scale (RnS): Choose a fixed angle θ and rotate the signal $f^d(x, y)$ by this angle. Perform scale normalization.
4. OII Computation: Compute the OII at the rotated position.
5. Iteration: Repeat steps 3 and 4 for a full rotation interval $\theta = [0, 2\pi]$.
6. Normalization: Select the signal corresponding to the maximum value of the OII and store it as the normalized signal f^n.
7. AIWT: Compute the wavelet transform of the normalized signal f^n.

The next section presents experimental results that demonstrate the BNA and the affine invariance of the AIWT.

4. EXPERIMENTAL RESULTS

We first choose the test signal f to be the binary image of a jet plane. We choose as the prototype signal f^o the output of the BNA applied to the test signal f. The result is shown in Fig.1(a). Fig.1(b) plots the orientation indicator index of the prototype signal as a function of the rotation angle $\theta = [0, 2\pi]$. Observe that the OII plot is periodic in $\pi/2$ as discussed earlier. For the test signal used in this experiment, there are four maximum points over the 2π interval of the OII plot. The orientation of the prototype f^o in Fig.1(a) corresponds to the maximum point at $\theta = 0$.

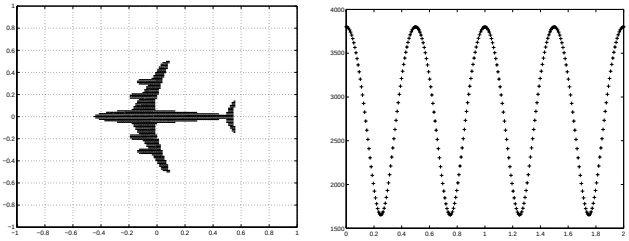
Fig.2(a) shows an affine distorted signal f^d obtained from the test signal f by applying the coordinate transform

$$\mathbf{A} = \begin{bmatrix} 1.6331 & -1.3745 \\ -1.5115 & 1.0507 \end{bmatrix}. \qquad (18)$$

The BNA is applied to f^d to recover the prototype signal f^o. The RnS step recovers the shape of the prototype signal from the affine distorted signal after three iterations in this experiment. The signal obtained at the third iteration of the RnS step is shown in Fig.2(b). The fixed rotation angle θ used in the RnS step is $3\pi/4$.

Fig.3(a) shows the OII plot of the distorted signal f^d. The OII plot of f^d after application of the BNA converges to a periodic function with distinct maxima. We choose the signal corresponding to the first maximum point and store it as the normalized signal f^n. The normalized signal f^n is a close approximation to the prototype signal f^o as shown in Fig.3(b). The normalized signal f^n is not an exact copy of the prototype signal f^o due to the discrete nature of the signals and the finite resolution of the OII plot.

The AIWT is now the wavelet transform of f^n. In this experiment, we use Daubechies 3 wavelets to compute the WT coefficients of the prototype, affine distorted, and normalized signals.



(a) Prototype Signal f^o (b) Orientation Indicator Index

Fig. 1. Prototype Signal f^o

Table 1 shows the maginitude, with its location index in the coefficient vector in parenthesis, of the three largest wavelet approximation and detail coefficients at the resolution level 3. We conclude that the AIWT coefficients are very close to the WT coefficients of the prototype signal occuring at the same spatial locations. In contrast, the WT coefficients of the distorted signal f^d differ significantly from the other two sets of coefficients as expected. The results are similar for other resolution levels and are omitted here.

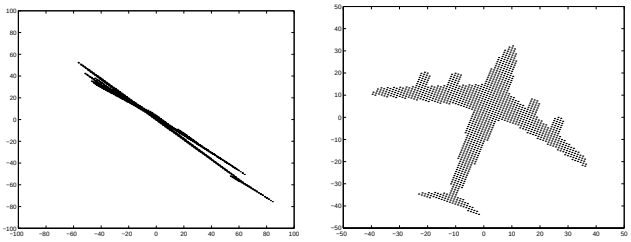
Fig.4 shows additional examples of the airplane test signal. Fig.5 shows the affine distorted versions of another test signal and the outcome of the BNA applied to them. These figures as well as other experiments with different distortions and test signals confirm the good performance of the BNA and the affine invariance of the AIWT.

5. CONCLUSION

In this paper, we presented the affine invariant wavelet transform. This transform is obtained by blindly normalizing the affine distorted signal replica to a prototype signal. The orientation of this normalized signal is determined by an orientation indicator index. Experimental results show a high level of invariance provided by the AIWT. The detailed analysis of the blind normalization algorithm and the orientation indicator index are to be presented elsewhere.

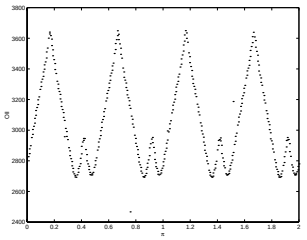
6. REFERENCES

[1] E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger, "Shiftable multiscale transforms," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 587–607, March 1992.

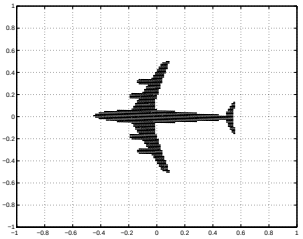


(a) Affine Distorted Signal f^d (b) Third Iteration of RnS Step

Fig. 2. The RnS Step



(a) OII of f^d

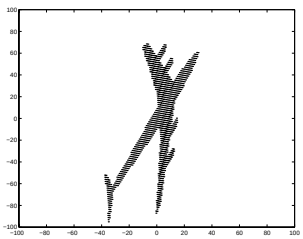


(b) Normalized Signal f^n

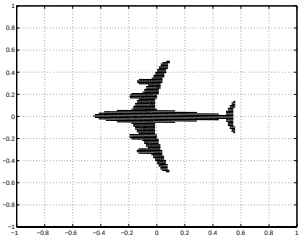
Fig. 3. The Blind Normalization Algorithm

Approx. Coefficients (index)			
WT of f^o	2155 (91)	2104 (119)	2208 (133)
AIWT of f^d	2120 (91)	2097 (119)	2204 (133)
WT of f^d	27.0 (91)	611.6 (119)	396.5 (133)
Horiz. Detail Coefficients (index)			
WT of f^o	609.6 (76)	598.1 (146)	-588.8 (147)
AIWT of f^d	606.8 (76)	584.8 (146)	-586.8 (147)
WT of f^d	4.2 (76)	444.1 (146)	-212.3 (147)
Vert. Detail Coefficients (index)			
WT of f^o	808.2 (88)	1009.9 (89)	821.1 (94)
AIWT of f^d	794.3 (88)	1010.3 (89)	829.2 (94)
WT of f^d	-0.2 (88)	-3.0 (89)	10.3 (94)
Diag. Detail Coefficients (index)			
WT of f^o	-241.0 (34)	260.6 (131)	-266.7 (161)
AIWT of f^d	-241.0 (34)	272.9 (131)	-266.3 (161)
WT of f^d	0.0 (34)	-314.6 (131)	1.8 (161)

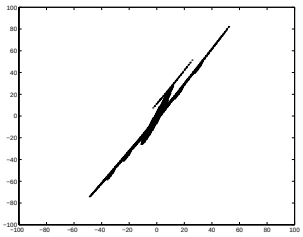
Table 1. The Wavelet Coefficients



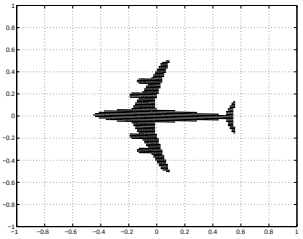
(a) Affine Distorted Signal f_1^d



(b) Normalized Signal f_1^n

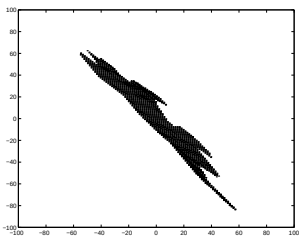


(c) Affine Distorted Signal f_2^d

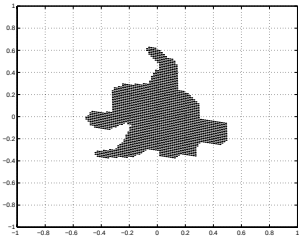


(d) Normalized Signal f_2^n

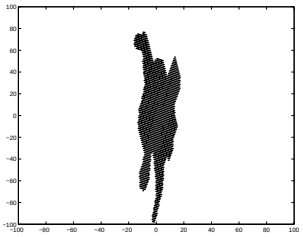
Fig. 4. Additional Examples



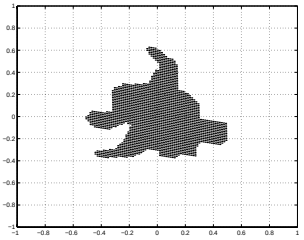
(a) Affine Distorted Signal f_3^d



(b) Normalized Signal f_3^n



(c) Affine Distorted Signal f_4^d



(d) Normalized Signal f_4^n

Fig. 5. Additional Examples

[2] A. Siebert, "A linear shift invariant multiscale transform," in *1998 International Conference on Image Processing*, 1998, vol. 3, pp. 688–691.

[3] Y. Hui, W. Kok, and T. Nguyen, "Theory and design of shift invariant filter banks and wavelets," in *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, 1996, pp. 49–52.

[4] I. Cohen, S. Raz, and D. Malah, "Shift invariant wavelet packet bases," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, 1995, vol. 2, pp. 1081–1084.

[5] S. A. Benno and J. M. F. Moura, "Scaling functions robust to translations," *IEEE Transactions on Signal Processing*, vol. 46, no. 12, pp. 3269–3281, December 1998.

[6] S. A. Benno, *Robust Representations*, Ph.D. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, December 1995.

[7] O. Rashkovskiy, L. Sadovnik, and N. Caviris, "Scale, rotation, and shift invariant wavelet transform," in *Proceedings of the SPIE*, 1994, vol. 2237, pp. 390–401.