行政院國家科學委員會專題研究計畫　成果報告

數位影像變形及內插處理器之研究

計畫類別：個別型計畫
計畫編號：NSC93-2622-E-002-013-CC3
執行期間：93 年 05 月 01 日至 94 年 04 月 30 日
執行單位：國立臺灣大學電機工程學系暨研究所

計畫主持人：貝蘇章

報告類型：精簡報告

處理方式：本計畫為提升產業技術及人才培育研究計畫，不提供公開查詢

中　華　民　國 94 年 5 月 11 日

數位影像變形及內插處理器之研究
Research on Digital Image Warping and Interpolation

計畫編號: NSC 93-2622-E-002-013-CC3
執行期限：93年8月1日至94年7月31日

主持人：貝蘇章 台灣大學電機系教授

摘要

本研究提出一任意尺寸大小的矩陣(Scaling matrices)，可以對數位影像做任意放大、縮小、變形、及內插，並可導出二度空間不均勻完美重建數位濾波器組，提供影像壓縮及傳輸方面的應用。

Abstract

In this paper, we propose a new method to realize an irrational scaling of a 2D digital signal/image. This method is suitable for any scaling matrices with real entries, including irrational numbers. Based on this algorithm, a new structure of a 2D four-channel non-uniform perfect reconstruction filter bank is derived. Through the proposed structure, it is now possible to segment the 2D frequency components of a 2D digital signal/image arbitrarily through a careful design of the scaling matrix while still preserving the PR property. Therefore, we can find many applications such as image compression and communication using this structure. The experimental results are presented at the end of this paper.

index terms—irrational scaling, interpolation, decimation, 2D nonuniform filter bank, perfect reconstruction

I. INTRODUCTION

In a conventional filter bank system, the scaling operation is very important in that it controls the sampling rate to meet the system requirements. As for the 2D filter bank system, the scaling matrix not only controls the sampling rate through its determinant but also incorporates shearing and rotating operations as those fundamental properties of a 2D affine transform. However, as long as a discrete time system is considered, most previous works fail to realize a scaling operation with irrational matrices. Although rational scaling can approximate any irrational scaling to reasonable accuracy in the 1D case, i.e. $r = M/N$ [1], M and N might have to be very large with higher sampling rates. Furthermore, as far as the 2D case is concerned, the decomposition of the scaling matrix to that with all integer entries is trivial and not practical.

In this paper, we propose a new method to realize any irrational scaling operations of a 2D digital signal. This method well preserves the linearity property while performing the scaling operation. In addition, the formula has a concise closed form in time domain, which is invulnerable to any computational approximation. Therefore, a 2D PR filter bank structure with irrational down-sampling matrices is realizable using the proposed scaling algorithms.

The paper is organized as follows. Section **II** provides the derivation of the new scaling algorithm. Section **III** introduces a decomposition method of the scaling matrix and the corresponding structure of the 2D irrational scaling PR filter bank is shown in Section **IV**. Section **V** presents the simulation results of the above scheme. Finally the conclusion is made in Section **VI**.

II. PROPOSED IRRATIONAL SCALING ALGORITHM

Given a 2D discrete time signal $x[m,n]$ with DTFT $X\left(e^{jw_m}, e^{jw_n}\right)$, suppose its scaled version is $x_A[m,n]$ with DTFT $X_A\left(e^{jw_m}, e^{jw_n}\right)$, where A is the scaling matrix and

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad a,b,c,d \in \Re \tag{1}$$

$$D = \left|\det(A)\right| = \left|ad - bc\right| \tag{2}$$

$$x_A\begin{bmatrix} m \\ n \end{bmatrix} = x\left[A\begin{bmatrix} m \\ n \end{bmatrix}\right] = x\left[\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} m \\ n \end{bmatrix}\right] \tag{3}$$

For simplicity, we will use $X(w_m, w_n)$, $X_A(w_m, w_n)$ instead of $X\left(e^{jw_m}, e^{jw_n}\right)$ and $X_A\left(e^{jw_m}, e^{jw_n}\right)$ respectively through out this paper. Since Eq.(3) is defined only for those A's with integer entries, we will seek for another solution using spectral approaches as follows. Referring to the scaling property of 2D continuous time Fourier transform (CTFT) as below [2],

$$x(u,v) \xleftrightarrow{CTFT} X(jw_u, jw_v)$$

$$x_A(u,v) \xleftrightarrow{CTFT} = \frac{1}{D} X\left(j\frac{dw_u - bw_v}{D}, j\frac{-cw_u + aw_v}{D}\right) \tag{4}$$

we could reasonably infer Eq.(4) to its DTFT counterpart, i.e.

$$x[m,n] \xleftrightarrow{DTFT} X(w_m, w_n)$$

$$x_A[m,n] \xleftrightarrow{DTFT} \frac{1}{D} X\left(\frac{dw_m - bw_n}{D}, \frac{-cw_m + aw_n}{D}\right) \tag{5}$$

Note that the above form is the corresponding discrete time

scaling formula which reveals that scaling by A in time domain is equivalent to scaling by $(A^{-1})^T$ in frequency domain and magnifying by $1/|\det(A)|$. Since $X(w_m, w_n)$ is a 2D continuous time periodic function with period $2p$ along both vertical and horizontal directions, $X_A(w_m, w_n)$ by definition exists whenever $a, b, c, d \in \Re$. Based on Eq.(5), we can now proceed to derive the 2D irrational scaling algorithm step by step.

$$X_A(w_m, w_n) = \frac{1}{D} X\left(\frac{dw_m - bw_n}{D}, \frac{-cw_m + aw_n}{D}\right)$$
$$= \frac{1}{D} \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m', n'] e^{-j\frac{dw_m - bw_n}{D}m'} e^{-j\frac{-cw_m + aw_n}{D}n'} \quad (6)$$

By Eq.(6),

$$x_A[m,n] = IDTFT\{X_A(w_m, w_n)\}$$
$$= \frac{1}{2p}\frac{1}{2p} \int_{-p}^{p}\int_{-p}^{p} X_A(w_m, w_n)\ e^{jw_m m}\ e^{jw_n n}\, dw_m dw_n$$
$$= \frac{1}{2p}\frac{1}{2p} \int_{-p}^{p}\int_{-p}^{p} \frac{1}{D} \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m', n']$$
$$e^{-j\frac{dw_m - bw_n}{D}m'} e^{-j\frac{-cw_m + aw_n}{D}n'} e^{jw_m m} e^{jw_n n} dw_m dw_n$$
$$= \frac{1}{D} \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m', n'] \mathrm{sinc}\left(m - \frac{dm' - cn'}{D}\right) \mathrm{sinc}\left(n - \frac{an' - bm'}{D}\right) \quad (7)$$

Eq.(7) can serve as an initial solution for 2D irrational scaling problems. However, it suffers from possible aliasing and imaging effects which are introduced by the scaling matrix A. Therefore, we will modify Eq.(7) according to two different kinds of scaling matrix in the next section.

III. DECOMPOSITION OF THE SCALING MATRIX

Any scaling matrix A could be decomposed into the product of two matrices, i.e. X and Y, as follows.

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ b & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ f & 1 \end{bmatrix}$$
$$= \begin{bmatrix} a & ah \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ bf & b \end{bmatrix} = XY \quad (8)$$

where

$$\begin{cases} a = a - \dfrac{bc}{d} & , \quad x-direction\ scaling\ \ factor \\ b = d & , \quad y-direction\ scaling\ \ factor \\ h = \dfrac{c}{ad - bc} & , \quad x-direction\ shearing\ \ factor \\ f = \dfrac{b}{d} & , \quad y-direction\ shearing\ \ factor \end{cases} \quad (9)$$

Note that we have merged the x-direction scaling and shearing matrices into one single matrix X in Eq.(8), as well as Y. We now explain the main motivation and concepts. In Section **II**, we have referred that scaling by A in time domain is equivalent to scaling by $(A^{-1})^T$ in frequency domain and magnifying by $1/|\det(A)|$. Therefore, shearing in the x-direction in time domain is equivalent to shearing in the y-direction in frequency domain due to the matrix transpose operation. Although shearing operation does not influence the $2p$ periodic property in the frequency domain along the corresponding direction, it does break the period along its orthogonal direction. It is the main reason why we combine the scaling and shearing operations along the same direction in time domain since they both change the frequency period along the same direction in frequency domain. Note that if the frequency period is changed along only one single direction, it is equivalent to a 1D operation. Therefore, following the decomposition process, a single 2D scaling operation becomes two subsequent 1D scaling operations except for the associated shearing operations.

In this way, when the decimation operation, i.e. $a, b > 1$, is concerned, we need to preserve the high frequency components which are supposed to be filtered out in Eq.(7) in order to attain PR. This can be accomplished by using only one additional channel since we are now dealing with the scaling operation along either x-direction or y-direction one at a time. The details will be described in the next section.

On the other hand, when the interpolation operation, i.e. $a, b < 1$, is concerned, we have to remove those unwanted high frequency images which are introduced when scaling a periodic signal to a different new period, i.e. from $2p$ to $\{2pa, 2pb\}$ in our case. Therefore, the derivation of the up-sampling algorithm should be slightly modified from Eq.(7) such that the upper and lower bounds of the integrals are limited to $[-p/a, +p/a]$ and $[-p/d, +p/d]$. Changing the integral upper/lower bound is equivalent to applying an ideal low-pass filter after up-sampling. Moreover, since up-sampling by A is equivalent to down-sampling by A^{-1}, those A's in Eq.(7) are replaced by A^{-1}'s in the following derivation.

$$x_A[m,n] = IDTFT\{X_{A^{-1}}(w_m, w_n)\}$$
$$= \frac{1}{2p}\frac{1}{2p} \int_{-p/d}^{p/d}\int_{-p/a}^{p/a} X_{A^{-1}}(w_m, w_n)\, e^{jw_m m} e^{jw_n n} dw_m dw_n$$
$$= \frac{1}{2p}\frac{1}{2p} \int_{-p/d}^{p/d}\int_{-p/a}^{p/a} D \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m', n']$$
$$e^{-j(aw_m + bw_n)m'} e^{-j(cw_m + dw_n)n'} e^{jw_m m} e^{jw_n n} dw_m dw_n$$
$$= \frac{D}{ad} \sum_{m'=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} x[m', n'] \mathrm{sinc}\left(\frac{m - am' - cn'}{a}\right) \mathrm{sinc}\left(\frac{n - bm' - dn'}{d}\right) \quad (10)$$

In conclusion, we summarize Eq.(7) and Eq.(10) in Table **I**.

Table I . 2D irrational up-sampling and down-sampling formulas

A is either of the form $\begin{bmatrix} a & ah \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 1 & 0 \\ bf & b \end{bmatrix}$, $a, b > 1$
Up Sample by "A" $\frac{D}{ad}\sum_{m'=-\infty}^{\infty}\sum_{n'=-\infty}^{\infty} x[m',n']\,\mathrm{sinc}\left(\frac{m-am'-cn'}{a}\right)\mathrm{sinc}\left(\frac{n-bm'-dn'}{d}\right)$
Down Sample by "A" $\frac{1}{D}\sum_{m'=-\infty}^{\infty}\sum_{n'=-\infty}^{\infty} x[m',n']\,\mathrm{sinc}\left(m-\frac{dm'-cn'}{D}\right)\mathrm{sinc}\left(n-\frac{an'-bm'}{D}\right)$

IV. 2D IRRATIONAL SCALING PERFECT RECONSTRUCTION FILTER BANK STRUCTURE

The fundamental structure of the 2D four-channel non-uniform PR filter bank using irrational scaling matrices is shown in Fig.1, where

$$\begin{cases} X' = \begin{bmatrix} a' & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & h \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a' & a'h \\ 0 & 1 \end{bmatrix} \\ Y' = \begin{bmatrix} 1 & 0 \\ 0 & b' \end{bmatrix}\begin{bmatrix} 1 & 0 \\ f & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ b'f & b' \end{bmatrix} \end{cases} \quad (11)$$

$$\frac{1}{a}+\frac{1}{a'}=1, \frac{1}{b}+\frac{1}{b'}=1 \quad (12)$$

Note that there is no need to design the pre-filters and post-filters in our PR FB structure since those filters have already been merged into the down-sampling and up-sampling kernels respectively in Table I , which could be revealed from the derivations in the previous section. The equivalent pre-filter we address is the ideal low pass filter whose cutoff frequencies locate at $w_{cx} = w_{cy} = \pm p$ and the equivalent post-filter is the ideal low pass filter whose cutoff frequencies locate at $w_{cx} = \pm p/a, w_{cy} = \pm p/d$. The basic goal of proposed structure is to perfectly reconstruct the output such that $y[m,n] = x[m,n]$ under any circumstances, even if a, b, c, d are irrational numbers. The up-sampling and down-sampling kernel have already been summarized in Table I while we now rearrange the formulas of X , X', Y and Y' in Table II .

Table II . Four kernel matrices X , X', Y and Y'

X	X'	Y	Y'
$\begin{bmatrix} a & ah \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} a' & a'h \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ bf & b \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ b'f & b' \end{bmatrix}$

V. EXPERIMENTAL RESULTS

To test the proposed PR FB structure in Fig.1, we take a 100x100 gray level LENA with a black surrounding outer ring as our input image $x[m,n]$. This dark region is adopted in order to avoid possible exceeding of the scaled image beyond the border which might cause uncomfortable visual defects. In fact, the proposed PR FB structure is suitable for any kinds of 2D signals, not only for images. Here the down-sampling matrix A is chosen as

$$A = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{bmatrix}$$

Since the index m', n' in Table I has a range from minus infinity to infinity which is not feasible in practical implementations, here we choose $m', n' = [-50, +50]$ in our experiment and observe the reconstruction errors due to truncation of m', n'. After reconstruction, the mean absolute value of the reconstruction error is only 0.2892, which is showed in Fig.2(g) together with the corresponding magnified version in Fig.2(h). Note that these errors are introduced by truncation of m', n', which makes the equivalent low-pass filters no ideal any more. However, the errors keep decreasing as m', n' grows larger and the filter bank structure virtually becomes perfect reconstruction as $m', n' \to \infty$.

VI. CONCLUSION

In this paper, a 2D non-uniform perfect reconstruction filter bank using irrational scaling matrices is proposed. The advantage of the new scaling algorithm is that it is a concise closed form formula in time domain, and it works well with any arbitrarily designed scaling matrices with irrational entries while still maintain the PR property. This provides us a powerful tool to extract any linear shapes of the 2D frequency components for subsequent applications, such as image/video compression and communication. The experimental results have shown the perfect reconstruction after a several irrational scaling operations.

APPENDIX

We include the detailed processed images and their corresponding frequency section diagrams in Fig.3 for reference. It is a step by step illustration of the entire structure in Fig.1. Note that the colors represent only the sections of the frequency domain, not the true values of it.

REFERENCES

[1] J. W. Adams, R. W. Bayma, and J. E. Nelson, "Digital Filter Design for Generalized Interpolation," *IEEE International Symposium on Circuits and Systems*, May 1989, Page(s): 1299 -1302 vol.2

[2] R. N. Bracewell, K. -Y. Chang, A. K. Jha, and Y. -H. Wang, "Affine theorem for two-dimensional Fourier transform," *Electronics Letters*, Volume: 29 , Issue: 3 , 4 Feb. 1993, Pages:304

[3] Seungsin Lee, W. Zhao, R. Narasimha and R.M. Rao, "Discrete-time models for statistically self-similar signals," *IEEE Transactions on Signal Processing*, Volume: 51, Issue: 5, May 2003, Page(s): 1221 -1230

[4] T. A. Ramstad, "Digital methods for conversion between arbitrary sampling frequencies," *IEEE Trans. on Acoustics, Speed, and Signal Processing*, Vol. ASSP-32, No. 3, June 1984, Page(s): 577 -591

[5] W. Zhao and R.M. Rao, "Discrete-time continuous-dilation wavelet transforms," *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, Oct 1998, Page(s): 233 -236

[6] W. Zhao, R.M. Rao, **"**Continuous-dilation discrete-time self-similar signals and linear scale-invariant systems," *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Volume: 3, Page(s): 1549 –1552

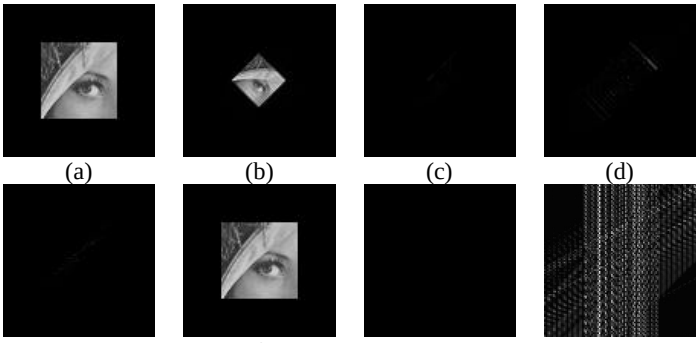


Fig.2 (a) Original input image. (b)-(e) Down-sampled images from channel 1 to channel 4 with down-sampling matrices XY, X'Y, XY', and X'Y' respectively. (f) Perfectly reconstructed output image. (g) Reconstruction error. (h) A magnified version of (g)

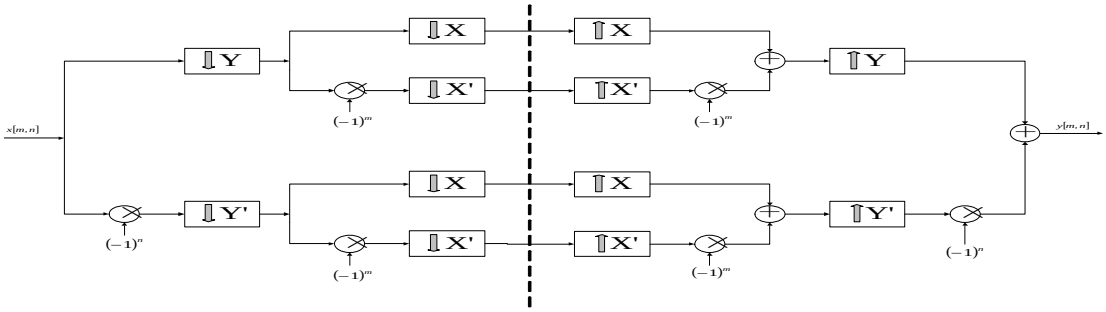


Fig.1 2D irrational scaling four-channel perfect reconstruction filter bank structure

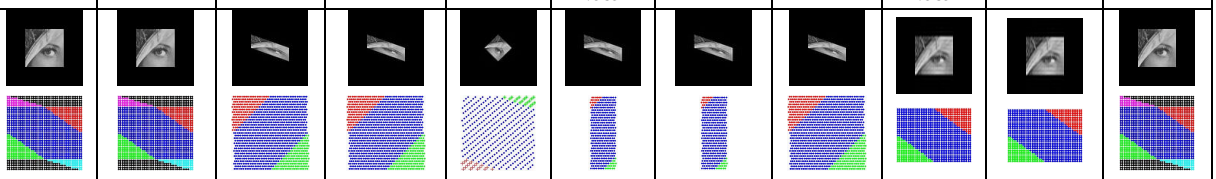
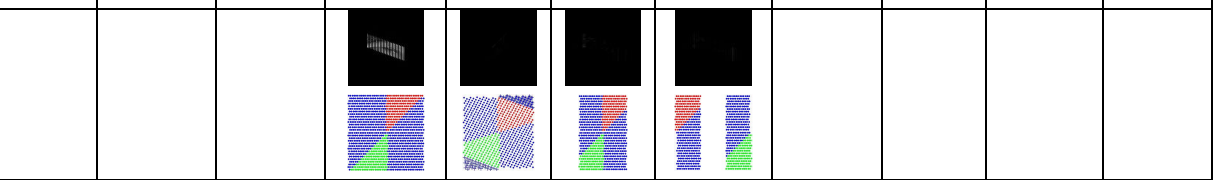
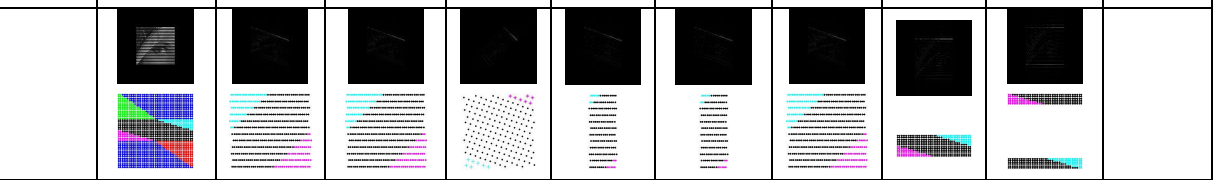
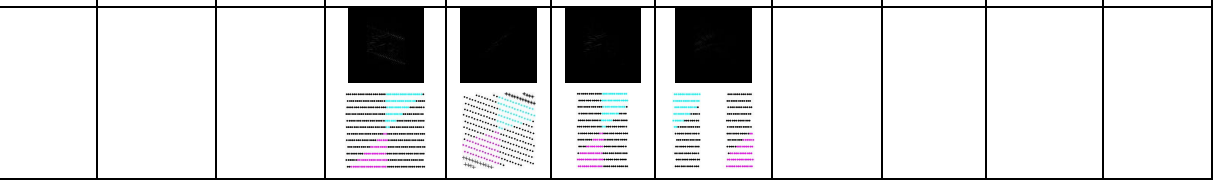
	original	modulation	down sample	modulation	down sample	up sample	demodulation	combination	up sample	demodulation	combination
			Y		X	inverse X			inverse Y		
Channel 01											
				horizontal	X'	inverse X'	horizontal				
Channel 02											
		vertical	Y'		X	inverse X			inverse Y'	vertical	
Channel 03											
				horizontal	X'	inverse X'	horizontal				
Channel 04											

Fig.3 Step by step illustration of Fig.1 in both spatial domain and frequency domain