International Journal of Computer Applications (0975 – 8887)
Volume 56– No.3, October 2012

Super Resolution Transcoding Algorithm in DCT domain using DFT Domain

Ankit Garg
Amity University
Manesar-Haryana, India

Khalid Hussain
Amity University
Manesar-Haryana, India

ABSTRACT

In this paper we have presented a transcoding algorithm to perform super-resolution of sub-sampled images. First of all we used 1D case in the fourier domain (DFT). Then we extended the same approach for the 2D case. After presenting the results for this we looked at the possibility of improving the performance of our algorithm. This was done by removing the need to perform matrix inversions (highly computation expensive operation). To map the operation to the DCT domain, we began by exploring the relationship between the DFT coefficients of a sequence with the DCT coefficients. Once the relationships were established we were able to extend our DFT approach to the DCT domain as well.

General Terms

DFT, DCT, MPEG

1. INTRODUCTION

With the expansion of digital media, digital images and videos are widely available for use and editing. Video compression algorithms are being used to compress digital video for a wide variety of applications, including video delivery over the internet, advanced television broadcasting, as well as video storage and editing. The performance of modern compression algorithms such as MPEG is quite impressive -- raw video data rates often can be reduced by factors of 15-80 without considerable loss in reconstructed video quality. However, the use of these compression algorithms often makes other processing tasks quite difficult. For example, many operations once considered simple, such as splicing and downscaling, are much more complicated when applied to compressed video streams. The goal of transcoding is to process one standards-compliant video stream into another standards-compliant video stream that has properties better suited for a particular application. This is useful for a number of applications. For example, a video server transmitting video over the internet may be restricted by stringent bandwidth requirements. In this scenario, a high-quality compressed bit-stream may need to be transcoded to a lower-rate compressed bit-stream prior to transmission; this can be achieved by lowering the spatial or temporal resolution of the video or by re-quantizing the MPEG data. Another important problem that arises in visual communications is the need to create an enhanced-resolution video image sequence from a lower resolution input video stream.

There are a number of methods for creating high-quality video or images from a lower-quality video. This can be done by either increasing the frame rate (by inserting number of frames in between two frames), called Temporal domain or be improving the frame resolution (by inserting more pixel points in the given frame), called Spatial domain .The latter involves prediction using the information of adjacent frames and then motion compensating a number of video frames to produce the desired video. These methods are formulated in space domain and require the input to be expressed in that format. We propose a motion-compensated transform-domain super-resolution procedure for creating high-quality video that directly incorporates the transform-domain quantization information by working with compressed bit stream [1], [2].

2. PROPOSED METHODOLOGY

We propose a motion-compensated transform-domain super-resolution procedure for creating high-quality video [3] that directly incorporates the transform-domain quantization information by working with compressed bit stream.

3. PROBLEM DEFINITION

Given: M frames each with sampling frequency $F = (F_x, F_y)$

To Generate: 1-Super resolution frame with sampling frequency $F' = (M_1F_x, M_2F_y)$

To define super–resolution more precisely, let's consider 1D case. Let x[n] be the given original sequence. Now we down-sample it by a factor of M and get M sub-sampled sequences. The k^{th} such sub-sampled sequence, $y_k[n]$, is defined as

$y_k[n] = x[Mn+k]$ where k =0,1,....,M-1

So the problem states that we are given these $y_k[n]$'s and we have to reconstruct the super-resolution sequence, x[n], back from these sub-sampled sequences.

4. INITIAL APPROACH

Let x(t) be a continuous-time signal that is sampled uniformly at t = nT, generating the sequence x[n] where

$$x[n] = x(nT), \qquad -\infty < n < \infty$$

where T being the *sampling period*. Now, the frequency-domain representation of x(t) is given by its continuous-time Fourier Transform (CTFT) X (jΩ),

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\Omega t}\, dt$$

whereas the frequency-domain representation of x[n] is given by its discrete-time Fourier transform (DTFT) $X(e^{j\omega})$,

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}$$

It can be shown very easily that the relationship between X(jΩ) and $X(e^{j\omega})$ is given by

$$X(e^{j\omega})=\frac{1}{T}\sum_{k=-\infty}^{\infty} X(\tfrac{j\omega}{T}-\tfrac{j2\pi k}{T})$$

Now if the continuous-time signal x(t) is sampled with down-sampling factor M i.e. is sampled uniformly at t = nT'= nMT, then generated sequence y[n] will be y[n] = x(nMT) = x[nM], $-\infty < n < \infty$.If $Y(e^{j\omega})$ is the discrete-time Fourier Transform(DTFT) of sequence y[n] then

Putting k = i + p*M where i = 0, 1 .. M-1 and $-\infty < p < \infty$

$$or \quad Y(e^{j\omega})=\frac{1}{M}\sum_{i=0}^{M-1} X\left(e^{\frac{j(\omega-2i\pi)}{M}}\right)$$

Now if sub-sampled sequences are $y_k[n]$ = x[Mn + k] i.e first shift the initial sequence x[n] and then sample it where k = 0,

$$Y(e^{j\omega})=\frac{1}{T'}\sum_{k=-\infty}^{\infty} X(\tfrac{j\omega}{T'}-\tfrac{j2\pi k}{T'}) \quad where\ T'=MT$$

1… M-1, then we have

So the problem statement is given $Y_k(e^{j\omega})$ values, we are trying to find out $X(e^{j\omega})$ which when converting to x[n] will give the resolution M times increased.

$$Y(e^{j\omega})=\frac{1}{M}\sum_{i=0}^{M-1}\frac{1}{T}\sum_{p=-\infty}^{\infty} X(\tfrac{j(\omega-2\pi i)}{MT}-\tfrac{j2\pi p}{T})$$

5. RELATION WITH DISCRETE FOURIER TRANSFORM (DFT)

5.1 One-dimension case

In case of finite-length sequences x[n] , n = 0,1,…, MN-1

only MN values of $X(e^{j\omega})$, called the *frequency samples*, at MN distinct points ,$\omega = \omega_k$, k = 0,1, .., MN-1 are called Discrete Fourier Transform DFT.

$$X(l)=X(e^{j\omega})\Big|_{\omega=\frac{2\pi l}{MN}}=\sum_{n=0}^{MN-1} x[n]e^{\frac{-j2\pi ln}{MN}} \qquad 0\le l\le MN-1$$

Now the sub-sampled sequences $y_k[n]$ will have N points where k = 0, 1, .., M-1

$$Y_k(l)=\frac{1}{M}\sum_{i=0}^{M-1} e^{\frac{jk2\pi(l-iN)}{MN}} X(l-iN) \qquad 0\le l\le N-1;\ \ 0\le k\le M-1$$

or in matrix form $X_{Mx1}(l)=A^{-1}_{MxM}(l)\ Y_{Mx1}(l) \quad 0\le l\le N-1$

Thus we have MN equations in MN variables. Solving which will give X[l] at MN equally spaced frequencies. Using the inverse Discrete Fourier transform (IDFT),

$$x[n]=\frac{1}{MN}\sum_{l=0}^{MN-1} X(l)e^{\frac{j2\pi ln}{MN}} \qquad 0\le l\le MN-1;\quad 0\le n\le MN-1$$

we will get x[n] with MN (M times more) points.

5.2 Extension to two-dimension

Now suppose the 2-D sequence is $x[n_1,n_2]$ and the sub-sampled sequences are $y_{k1k2}[n] = x[M_1n_1 + k_1, M_2n_2 + k_2]$ where $k_1 = 0,1 \ldots , M_1-1$; $k_2 = 0,1 \ldots , M_2-1$.Then the relation between DFT's of $y_{k1k2}[n]$ and $x[n_1,n_2]$ is given by

$$\boldsymbol{Y_{k_1k_2}(l,m)=\frac{1}{M_1M_2}\sum_{i=0}^{M_1-1}\sum_{j=0}^{M_2-1} e^{\frac{jk_12\pi(l-iN_1)}{M_1N_1}} e^{\frac{jk_{12}2\pi(l-iN_1)}{M_1N2}} X(l-iN_1, m-jN_2)}$$

where $0 \le l \le N_1 - 1$ and $0 \le m \le N_2 - 1$;
$0 \le k_1 \le M_1 - 1$ and $0 \le k_2 \le M_2 - 1$

Once again get X(l,m) at (M_1N_1, M_2N_2) frequency points. Calculate IDFT to get $x[n_1, n_2]$ with (M_1N_1, M_2N_2) points in space and hence with increased resolution.

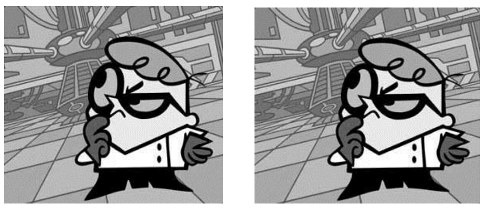
5.3 Experimental results

The algorithms we described above were experimented upon an image. The results are shown below.

Fig 1: Original Image

Fig 2: Sub-sampled images; (a) $k_1 = 0, k_2 = 0$, (b) $k_1 = 0, k_2 = 1$, (c) $k_1 = 1, k_2 = 0$, (d) $k_1= 1, k_2 = 1$

5.4 Image reconstruction

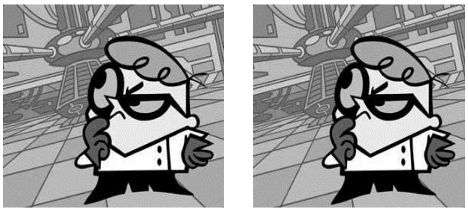


(a) (b)

Fig 3: Image Reconstruction using DFT; (a) Original image, (b) Reconstructed image

5.5 Comparison with interpolation

If instead of Super-Resolution, we do linear interpolation [4] i.e. insert the average values of pixel values in between two pixel both in horizontal direction as well as vertical direction, we get the following result,



(a) (b)

Fig4: Comparison with Interpolation; (a) Super-Resolution image, (b) Interpolated image

Clearly the one achieved from interpolation is blurred as compared to that from super-resolution

5.6 Improvement in DFT approach

The algorithm we purposed initially using DFT equation

$$X_{Mx1}(l)=A^{-1}_{MxM}(l)\ Y_{Mx1}(l) \qquad 0\le l\le N-1$$

involves matrix inversion. So we suggest a new approach to the problem. This can be understood by an example. Suppose the original sequence x[n] was sub-sampled to two sequences $y_0[n]$ and $y_1[n]$ having N points (Fig (a) and (b)). Now to get the sequence x[n] back, we need to do the following operations.

a) Insert M-1(here M = 2) zeros between every two consecutive samples of $y_0[n]$ Fig (c)). So this new sequence $y_{0ext}[n]$ has MN points.
b) Insert M-1 zeros between any two consecutive samples of $y_1[n]$ and then do circular shifting by one (Fig (d)). Again the new sequence $y_{1ext}[n]$ has MN points.
c) Add the two sequences to get x[n] back (Fig (e)).

The method was shown in time-domain. In DFT domain, the insertion of zeros means periodic extension and circular shifting means multiplication with exponential. So the relation between DFT's of $y_k[n]$ and x[n] is given by

$$X_{MN}(l)=\sum_{k=0}^{M-1} Y_k(<l>_N)\ e^{\frac{-j2\pi kl}{MN}} \qquad 0\le l\le MN-1$$

which can be easily extended to 2-D as

$$X_{(M_1N_1,M_2N_2)}(l,m)=\sum_{k_1=0}^{M_1-1}\sum_{k_2=0}^{M_2-1} Y_{k_1k_2}(<l>_{N_1},<m>_{N_2})\ e^{\frac{-j2\pi k_1 l}{M_1N_1}}\ e^{\frac{-j2\pi k_2 m}{M_1N_2}}$$

$$0\le l\le M_1N_1-1; \qquad 0\le m\le M_2N_2-1$$

Using these relationships we implemented the algorithms for both 1-D and 2-D DFT's.

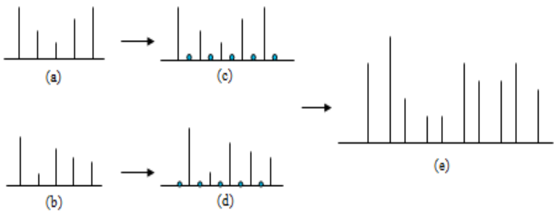


Fig 5: Improved Approach (a) and (b) are sub-sampled images $y_0[n]$, $y_1[n]$; (c) and (d) are extended sequences $y_{0ext}[n]$, $y_{1ext}[n]$; (e) is re-constructed sequence x[n]

5.7 Super-resolution using discrete cosine transform (DCT)

As we are going to work in transformed domain only and MPEG frames use DCT (Discrete Cosine Transform) domain for compression, we have to do this super-resolution in DCT domain. But there are no direct relationships in DCT domain, as we have in DFT domain, corresponding to the time domain operations like,

1) Shifting
2) Down-sampling with shifting

So our next task is to establish relationships between DFT and DCT of two sequences, which will enable us to do the super-resolution of the given sub-sampled images with DCT coefficients only.

5.8 Relationships between DFT and DCT

5.8.1 One-dimension case

Given an N-length sequence
x[n] = {x[0], x[1],….,x[N-2], x[N-1]}
Its N-point DCT is given by

where **α(k)** = √(1/N) for k =0
√(2/N) else

Now consider a 2N length symmetric sequence given by
y[n] = {x[N-1], x[N-2],…..,x[1], x[0], x[0], x[1],….,x[N-2], x[N-1]}
If Y(k) is 2N- point DFT of y[n] then it can be shown that

5.8.2 Extension to two-dimension

$$where \quad \beta(k)=\frac{2e^{\frac{-j\pi k(2N-1)}{2N}}}{\alpha(k)} \qquad 0\le k\le N-1$$

Now consider a 2-D sequence x[m,n], m = 0,1…N_1-1 ; n = 0,1…N_2-1; Its 2-D DCT is given by

$$C(k,l)=\propto(k,l)\sum_{m=0}^{N_1-1}\sum_{n=0}^{N_2-1} x[m,n]\cos\frac{(2m+1)\pi k}{2N_1}$$

$$\cos\frac{(2n+1)\pi k}{2N_2}\ 0\le k\le N_1-1; 0\le k\le N_1-1;$$

where **α(k,l)** = **α(k) * α(l)** and α(k) , α(l) are defined as above

.Following in the same manner as we did in 1D case, Consider symmetric extension sequence given by y[m,n] = x[N_1-1-m, N_2-1-n] for m = 0,1,..N_1-1 ; n = 0,1,..N_2-1x[N_1-1-m, n- N_2] for m = 0,1,..N_1-1 ; n = N_2, N_2+1,..2N_2-1x[m- N_1,N_2-1-n] for m = N_1, N_1+1,..2N_1-1 ; n = 0,1,..N_2-1 x[m- N_1,n- N_2] for m = N_1, N_1+1,..2N_1-1;n = N_2,N_2+1,2N_2-1

If Y(k,l) is 2D-DFT of y[m,n] then again it can be shown that

Y(k,l) = β(k,l)*C(k,l)

$$where \quad \beta(k,l)=\frac{4e^{\frac{-j\pi k(2N_1-1)}{2N_1}}\,e^{\frac{-j\pi l(2N_2-1)}{2N_2}}}{\alpha(k,l)} \quad 0\le k\le N_1-1;\ 0\le l\le N_1-1$$

Now if we sample the sequence x[n] and make the corresponding symmetric sequence y[n], it has no relationship with earlier symmetric extended sequence.

x[n] = {1,2,3,4,5,6} ⟶ y[n]= {6,5,4,3,2,1,1,2,3,4,5,6}

x_{samp} = {1,3,5} ⟶ y_{samp}[n] = {5,3,1,1,3,5}

So we can't work with symmetric extended sequence y[n] !!

x[n] = {1,2,3,4,5,6} ⟶ y_l[n]= {6,5,4,3,2,1,0,0,0,0,0,0}

x_{samp} ={1,3,5} ⟶ y_{lsamp}[n]={5,3,1,0,0,0}

But we can work with left sided sequence y_l[n] defined as,
y_l[n] = {x[N-1], x[N-2],…..,x[1], x[0],0, 0,....,0}
So the next task is to relate y[n] and y_l[n].
The relationships between Y_l(k) and Y(k) can be derived easily as,

$$a(k)(1+\cos\tfrac{\pi k}{N})+b(k)\sin(\tfrac{\pi k}{N})=a_0(k)\ ;0\le k\le N-1$$

$$\sum_{r=1}^{N-1}\{2a(N-r)\cos(\tfrac{\pi rn}{N})+2b(N-r)\sin(\tfrac{\pi rn}{N})\}+(-1)^n a(0)+a(N)=0$$

$$where \quad n=N,N+1,.....,2N-1$$

or in matrix form, $$x=A^{-1}b$$

where $Y_l(k) = a(k) + j\,b(k)$;$Y(k) = a_0(k) + j\,b_0(k)$

This method involves inversion of a matrix. So we suggest a better approach in the next section.

5.9 *Improvement in DCT approach*

5.9.1 One-dimension case

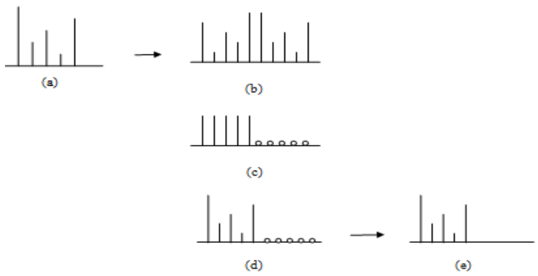


Fig 6: Time domain approach for DCT to DFT conversion. (a) original sequence x[n];(b) symmetrical extended sequence x_{ext}[n]; (c) windowing sequence w[n]; (d) left sided sequence y[n]; (e) output sequence y_0[n]

In time domain multiplication

y[n] = x[n] w[n]

where
$$w[n]=\begin{cases}1 & 0\le n\le N-1\\ 0 & N\le n\le 2N-1\end{cases}$$

is circular convolution in DFT domain defined as

$$Y(k)=\frac{1}{2N}\sum_{m=0}^{2N-1}X(m)\,W_N(<k-m>_N) \quad 0\le k\le 2N-1$$

where

$$W_N(k)=\begin{cases}N & k=0\\ 1-j\cot(\frac{\pi k}{2N}) & k=1,3,\ldots.,2N-1\\ 0 & k=2,4,\ldots,2N-2\end{cases}$$

5.9.2 Extension to two-dimension

Results can easily be extended to 2-D as

$$Y(k,l)=\frac{1}{4N_1N_2}\sum_{m=0}^{2N_1-1}W_{N_1}(<k-m>_{N_1})\sum_{n=0}^{2N_2-1}X(m,n)\,W_{N_2}(<l-n>_{N_2})$$
$$0\le k\le 2N_1-1;\ 0\le l\le 2N_2-1$$

where $W_{N1}(k)$ and $W_{N2}(k)$ have their usual meanings.

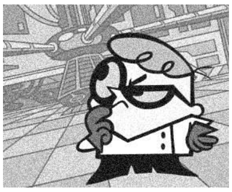
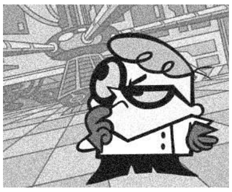
5.9.3 Experimental results

As we have already said that super-resolution directly in DCT domain is not possible. So we approached to the problem by first converting the DCT's of sub-sampled images to DFT of extended sequences. From there we got the super-resolution image in DFT domain using the earlier formulae. And then we converted back the DFT domain image to DCT domain which is super-resolution image of given DCT domain images.
The results for 2-D DCT case are shown below:

Fig 7: Original Image with 208 X 222

- Input image (Fig 1.7) was sampled at M_1=2, M_2=2
- Four sub-sampled images were generated (Fig 1.8)
- Using them Super- Resolution image is created


(a) (b)

(c) (d)

Fig 8: Sub-sampled images with 104 X 111; (a) $k_1 = 0$, $k_2 = 0$, (b) $k_1 = 0$, $k_2 = 1$, (c) $k_1 = 1$, $k_2 = 0$, (d) $k_1 = 1$, $k_2 = 1$

5.10 IMAGE RECONSTRUCTION


(a)

(b)

Fig 9: Image Reconstruction using DCT; (a) Original image, (b) Reconstructed image

6. CONCLUSION

When we are making a super-resolution frame form M frames, it may be possible that you don't have enough sub-samples of a particular object. For example you are making one super-resolution frame using the inter-dependency of three frames. Then suppose we find an object, which is present in only one of these frames. So for that object you have to resize it by any arbitrary ratio L/M for the new super-resolution frame. And once again we have to do it in the compressed domain i.e. DCT domain. So we have also implemented this feature in this paper.

In this paper we have presented a transcoding algorithm which is used in video processing and communication [5] to perform super-resolution of sub-sampled images. As a first, we looked at the 1D case in the fourier domain (DFT). Then we extended the same approach for the 2D case. After presenting the results for this we looked at the possibility of improving the performance of our algorithm. This was done by removing the need to perform matrix inversions (highly computation expensive operation). To map the operation to the DCT domain, we began by exploring the relationship between the DFT coefficients of a sequence with the DCT coefficients. Once the relationships were established we were able to extend our DFT approach to the DCT domain as well.

When we are making a super-resolution frame form M frames, it may be possible that you don't have enough sub-samples of a particular object. For example you are making one super-resolution frame using the inter-dependency of three frames. Then suppose you find an object, which is present in only one of these frames. So for that object you have to resize it by any arbitrary ratio L/M for the new super-resolution frame. And once again we have to do it in the compressed domain i.e. DCT domain. So we have also implemented this feature in this paper.

7. FUTURE WORK

We can implement an approach towards incorporating spatial filtering operations directly in the DCT domain. We can began with exploring the relationship between the DCT and DFT coefficients. The relationship used here is different from the one discussed in this paper and was aimed primarily at removing the redundancies that arise from double-side replication of the input sequence. We can present an approach to shift the input coefficients by a single sample, which can be generalized as shift by any amount can be looked upon as a series of one-sample shifts. We can present our extension to the 2D case on the basis of the argument that shifting operation is separable as well as the DCT is separable and hence the x and y shifts can be taken care of independently.

8. REFERENCES

[1] R.Y. Tsai and T.S. Huang, "Multiframe Image Restoration and Registration, Chapter 7", in *Advances in Computer Vision and Image Processing* (T.S. Huang, ed.), JAI Press Inc., 1984.

[2] Y. Altunbasak, A. J. Patti, and R. M. Mersereau, "Super-Resolution Still and Video Reconstruction from MPEG-Coded Video", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No. 4, April 2002.

[3] Altunbasak, Y.; Patti, A.J.; Mersereau, R.M. "Super-resolution still and video reconstruction from MPEG-coded video", IEEE Transactions on Circuits and Systems for Video Technology, Volume: 12, Issue: 4, 2002.

[4] Gilman, A.; Bailey, D.G.; Marsland, S.R. "Interpolation Models for Image Super-resolution", Electronic Design, Test and Applications, 2008, Publication Year: 2008.

[5] Yao Wang, Jorn Ostermann, and Ya-Qin Zhang. "Video Processing and Communication" Prentice Hall Signal Processing Series, pp. 406-454, 2002.

[6] Vetro, A.; Christopoulos, C.; Huifang Sun,"Video transcoding architectures and techniques:an overview" Signal Processing Magazine, IEEE Volume: 20 , Issue: 2 , 2003.

[7] HyunWook Park, *Senior Member IEEE*, YoungSeo Park and Seung-Kyun Oh, "L/M-fold Image Resizing in Block-DCT Domain using Symmetric Convolution", IEEE Transactions on Image Processing, Vol. 12, No. 9, Sep 2003