行政院國家科學委員會專題研究計畫　成果報告

子計畫四:MPEG-4 單晶片無線多媒體通訊系統

計畫類別： 整合型計畫
計畫編號： NSC91-2218-E-009-005-
執行期間： 91 年 08 月 01 日至 92 年 07 月 31 日
執行單位： 國立交通大學電子工程學系

計畫主持人： 蔣迪豪

計畫參與人員： 黃名彥、蘇子良、王士豪、王俊能

報告類型： 完整報告

處理方式： 本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中　華　民　國 92 年 10 月 30 日

行政院國家科學委員會補助專題研究計畫 ☑ **成 果 報 告**
□期中進度報告

單晶片無線多媒體資訊家電之設計與製作(3/3)
子計劃四 ： 單晶片無線多媒體通訊系統

計畫類別：□ 個別型計畫　　☑ 整合型計畫
計畫編號： NSC－91-2218-E-009-005
執行期間： 91 年 8 月 1 日至 92 年 7 月 31 日

計畫主持人：蔣迪豪 交通大學電子工程系所 副教授
共同主持人：
計畫參與人員： 黃名彥、蘇子良、王士豪、王俊能

成果報告類型(依經費核定清單規定繳交)：□精簡報告 ☑完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
☑涉及專利或其他智慧財產權，□一年 □二年後可公開查詢

執行單位：交通大學電子工程系所

中 華 民 國 92 年 10 月 30 日

行政院國家科學委員會專題研究計畫成果報告

單晶片無線多媒體資訊家電之設計與製作(3/3)
子計劃四 ： 單晶片無線多媒體通訊系統

計畫編號：91-2218-E-009-005
執行期限：91 年 8 月 1 日至 92 年 7 月 31 日
主持人：蔣迪豪　　國立交通大學電子工程學系 副教授

目錄 Table of Contents

中文摘要

本計畫從事基於 MPEG-4 視訊壓縮標準之壓縮與解壓縮器在抗噪性與錯誤回復力方面和即時性之研究。在壓縮器方面，我們提出一個 R-D(Rate-Distortion)最佳化的視訊內容自動 Intra 更新的方法。藉由傳輸頻道的錯誤對於視訊位元資料的影響之統計特性，我們可以使在解壓縮端有效地控制錯誤資料在可回復的視訊畫面之間傳遞間隔。在解壓縮器方面，我們針對 MPEG-4 視訊壓縮標準在網路傳輸過程中遇到的封包傳輸錯誤, 提出防止解壓縮器遇到封包漏失的處理方法，並以計算機與特定的通道錯誤模型來模擬檢視其效能。除此，為了改善影像品質，我們也對回復的影像給予簡單的錯誤補償。為了完成在 ARM 平台上即時性的壓縮與解壓縮器，我們也採用一些快速的演算法與高效率的記憶體存取方法來改善壓縮與解壓縮器執行效能，並與予實現計算複雜度高的硬體。結合以上技術以及效能最佳化之壓縮與解壓縮器，將提高 MPEG-4 壓縮與解壓縮器之應用於單晶片無線多媒體資訊家電。

關鍵詞：

MPEG-4，自動 Intra 更新，錯誤回復力，錯誤補償，快速(反)餘弦轉換，快速動態偵測，硬體實現。

Abstract

MPEG-4 video coding standard provides applications for both Internet and mobile links. This article describes a way to construct an error resilience system including encoder and decoder. In the encoder end, to prevent error propagation, adaptive intra refresh is used. Due to the rapid increase of intra coding bits, some redundancy may be filled in the coding bitstream with minor effects. To optimize this effect, rate-distortion optimized mode switch is used according to current channel condition. To prevent decoding crash, resynchronization marker is used to skip the corrupted bitstream. In the decoder end, hybrid error concealment is used to statically optimize the concealed images.

Toward SoC project, our MPEG-4 encoder and decoder are ported onto Linux platform that can be run on ARM-9 device. To enhance the computation ability for real-time application, hardware implementation is adopted on FPGA. Motion estimation for encoder and Inverse DCT for decoder are realized on the proposed ARM platform, and communicate with each other via ARM Master Bus Architecture (AMBA). With the acceleration, 42% and 15% speedup are achieved for encoder and decoder respectively.

Keywords:

MPEG-4, Error Resilience, Error Concealment, Adaptive Intra Refresh (AIR), Rate-distortion optimization, fast DCT/IDCT, fast motion estimation, hardware implementation.

一、背景與目的

此計畫的主要目標為設計一個強韌而抗錯的 MPEG-4 視訊編碼系統。MPEG-4 視訊編解碼系統將建立於 Linux 作業平台以及 SoC 環境上。本子計畫年度目標包括(一) MPEG-4 視訊編解碼系統在 Linux 作業系統上之模擬與設計(二) MPEG-4 視訊編解碼系統在無線傳輸上強韌抗錯之模擬與設計。

無線通訊的普及化是未來的趨勢。目前數位無線通訊的技術正在蓬勃的發展中。現在無線網路的承載內容大都以語音為主。然而在可預見的將來，數位多媒體無線通訊的服務將成為主流產品。本計畫把發展一套抗錯編碼技術分解為兩個問題包括:

(一) 探討在 Linux 作業系統下執行 MPEG-4 視訊編解碼: 如何在以 Embedded Linux 為基礎的環境下執行困難的視訊壓縮功能是技術上的一大挑戰。此時將假設無傳輸錯誤，主要在解決編解碼演算法軟體程式實現時之問題。

(二) 探討在無線傳輸時強韌抗錯編解碼: 如何在位元有錯誤時補償視訊品質是一項困難的技術。MPEG-4 提供了一些編解碼的工具但並未指定如何使用。我們希望發展出一套符合標準規格且強韌之解碼演算法。

本子計畫成果的應用價值在可組合數位無線通訊之新標準規格 IEEE 802.11 及多媒體通訊標準 MPEG-4 而產生一極具有競爭力之先進無線通訊產品。其學術價值亦可從最近相關論文不斷發表於國際會議中可看出端倪。

二、報告內容

1. Introduction

MPEG-4 video coding standard [1] is developed to provide users a new level of performance for various video communications services such as video on demand (VOD) over the Internet. For video streaming over the Internet, the bitstreams will be corrupted by random error or packet loss in the channels. Thus, it is a challenge to realize an error resilient MPEG-4 video system for multimedia applications over the Internet [2]-[6]. To recover the video contents from the corrupted bitstreams during transmission, the resilient decoding process and the robust bitstreams are required.

The error recovery methods utilize all the useful information available at the receiving end for resynchronization of decoding processes. There are basically three types of information including spatial-temporal information, duplicate header information and the psychovisual properties of video [2]. In various implementations of video decoder, there is a tradeoff between the amount of the information exactly used and the final visual quality for the error resilience and error concealment. Based on the spatial-temporal characteristics of transmitted video and redundant header information, the MPEG-4 standard has provided several error resilient tools [2]. These error resilient tools are used to resynchronize the coding processes, localize the errors, and conceal the errors. The error detection and concealment techniques are informative in the MPEG-4 video coding specification. Thus, advanced error detection and concealment methods were provided to improve the reconstructed video quality.

In [14], the authors use nearby MB information to interpolate the lost MB. Since the distance of 4 sides from the correct nearby MB is not always the same, weighted function is applied according to the distance to get the best interpolation results for spatial concealment. In [15], the authors adopt nearby motion vectors to represent the current lost MB motion vectors (MV), and do the temporal concealment. To find the best one from nearby candidates, they evaluate the boundary

pixels to find minimal mean square errors as the elected one. Both two kinds of methods have its own advantages for fast motion and slow motion sequence. To combine the advantages over the 2 kind of methods, the hybrid concealment method is proposed to further improve the reconstructed quality [14].

In addition, to increase reconstructed quality by bending the error drift, we inserting intra coding blocks into the proper locations in encoding. As the intra refresh techniques, one of the effective ways is to add intra blocks randomly in the encoded bitstream to prevent error drift, which is called intra refresh (IR). The random IR process may waste bits, since the error drift occurs only for inter frame prediction. To further improve IR, Cote et al. [13] proposed an adaptive IR using rate distortion constraint optimization, which could save bits and improve the robustness of the encoded bitstreams simultaneously for video content delivery. To optimize the recovered picture quality, the joint optimization through the intra blocks insertion and the error recovery is addressed in the proposed video system for the Internet applications.

In our streaming system of error resilience capability, the statistics of error occurrence probability is estimated with the spatial-temporal concealment methods and the assumed channel models as in [13]. The spatial-temporal concealment method is identical in the reconstruction routine of the encoder and the decoding processes, which can synchronize the picture quality of the sender and the receiver. With the probability distribution of greater precision, the encoder can properly select the intra coded blocks to improve the picture quality of reconstructed video with less overhead. The simulation results show improvement of reconstructed picture quality by 6.33-13.22 dB on PSNR when compared to that is available by the MPEG-4 reference software.

2. The Architecture of Error Resilience MPEG-4 System

The proposed MPEG-4 error resilience system is as depicted in Fig. 1 and Fig. 2. Fig. 1 shows the functional encoding blocks to generate error resilient bitstreams. The dash-line in Fig. 1 means the calculation of coding bits and distortion for both intra and inter macroblock (MB). For each MB type, the error occurrence probability models over the real Internet are computed by passing the bitstream through the network simulator. With the error occurrence models for each MB, the accumulated lost rate for the MB at the same position within each frame is used to calculate the rate-distortion (R-D) cost. The R-D cost is then used for an optimized mode decision to assign the handling MB as an inter block or an intra block. To further realistically get the situation in the decoder end, we model the error concealment that is identical to that in decoder and get the reconstructed information to calculate the R-D cost. Based on the R-D cost for each MB type, the optimized mode is decided for the current MB and the relative bits are transmitted over the Internet to provide a video sequence with acceptable quality.

Fig.2 shows the functional blocks for an error resilience decoder in our streaming system. The received bitstream is parsed to seek for continuous resynchronization markers (RM). The successful bitstream parsing indicates no syntactic errors happen, and then normal decoding process is applied. If there is any syntactical error, the decoder will jump to the next RM for resuming decoding processes. After one frame is fully reconstructed, the proposed error concealment algorithm is applied into the concealment process of the reconstructed image based on the available information from the received bits.

3. Rate-Distortion Optimized Intra Refresh

To enhance the ability of error resilience, we proposed relative solutions over both the encoder and decoder end. In the encoder end, rate distortion optimized intra-refresh is proposed to improve the bit-stream

structure according to the network condition.

3.1Rate distortion optimized intra-refresh

Rate distortion optimized intra-refresh (RDIR) has been proposed for solving error propagation more effectively [13]. Intra refresh technique is to insert intra coding block instead of inter coding block in P frame to prevent serious error propagation over error prone network. Since the intra coding block sacrifice more bits, it will become inefficient when network condition changes with time. To improve this, intra block insertion with rate-distortion optimization adaptive to channel condition can bring us the most compact and least overhead encoder system.

The RDIR design flow can be depicted as Fig. 3. For each macroblock (MB), we calculate the cost for intra and inter blocks said $J_{\text{int}ra}$ and $J_{\text{int}er}$ by following Lagrange formula.

$J = D_q + \lambda \cdot R$

J : Lagrangian cost

λ : Parameter used to control coding bit rate in encoding process

D_q : Distortion induced from residue quantization

R : Bits used in coding a macroblock

After the cost of J is decided, the mode with minimal value of J is chosen as the current MB coding mode. For error prone environment, the distortion of D will suffer more serious quality loss. These include not only the original quantization error, but also the errors introduced when concealing the lost MB from nearby MB. So the above formula needs to be modified as

$J = \left(D_q \cdot (1-p) + D_c \cdot p\right) + \lambda \cdot R$

D_q : Distortion induced from residue quantization

D_c : Distortion induced from no-so-perfect concealment algorithm

p : Channel packet loss rate

To achieve the R-D optimization under the proposed intra-refresh encoding, the parameter of λ needed to be updated every frame to control the used bits under the same distortion. The updating formula follows

$$\lambda_{n+1} = \lambda_n \left(1 + \alpha\left(\sum_{i=1}^{n} R_i - nR_{\text{target}}\right)\right) \quad , \alpha = \frac{1}{20 \cdot R_{\text{target}}}$$

The parameter of α is decided from a verity of experimental trial for buffer control. The packet loss rate is used to model the internet protocol. Using network condition to model the situation in decoder end is expected to reconstruct better image quality. And if the modeling is 100% matched, we will get the same quality as transmitted one in error prone environment.

4. Error Resilience on Decoder End

To improve the reconstructed video quality over error prone network, both improved

resynchronization scheme and error concealment techniques are implemented in the decoder end.

4.1 Error Robustness over Packet Loss

To handle with packet loss, resynchronization markers (RM) are enabled to stop the collapse of decoder. When meeting with non-continuous MBs in the decoding process, the decoder will skip to next resynchronization marker and restart decoding again. Since the middle parts from the error starting point to next RM will be dropped due to the uncertain content information, the length between RMs may have great influence over the reconstructed quality. If the length is long enough to be able to contain whole blocks information, it will suffer serious quality information loss when meeting packet loss. In another word, if the length is too short, the redundant information will be distributed in the bit-stream and make the bit-stream size grow up quickly without significant improvement. The tradeoff can be chosen according to the application domain. Here we choose 1000 bits as the length after considering the application of VOD application under the bit-rate of above 256k bits per second (bps).

4.2 Hybrid Error Concealment

Contrast to RM which is to stop error from propagation, error concealment is to further improve the reconstructed image quality from existing information. There are two basic kind of concealment algorithm including spatial and temporal concealment. Spatial concealment which reconstructs the error MB from neighbor MB of the same frame can achieve better performance in fast motion sequence. The reason is that it is hard to track the trajectory with fast motion. Temporal concealment conceals current error MB by using temporal correlation. To apply to slow motion sequence, it can obtain better performance without obvious blocky artifact introduced by spatial concealment. To tradeoff these two methods, we proposed an improved hybrid concealment method.

The hybrid concealment algorithm is depicted as Fig. 4. For temporal concealment, we find the best matched motion vectors (MV) from neighbor MBs as shown in Fig. 5. The criterion for choosing the best matched motion vector follows the Minimal Mean Square Error (MMSE) shown below.

$$MSE =$$

$$\sum_{j=0}^{15}[P(x, j) - P(x-1, j)]^2 + \sum_{j=0}^{15}[P(x+15, j) - P(x+16, j)]^2 +$$

$$\sum_{i=0}^{15}[P(x+i, j) - P(x+i, j-1)]^2 + \sum_{i=0}^{15}[P(x+i, j+16) - P(x+i, j+15)]^2$$

The pixels used for calculating MSE is shown in Fig 6. We evaluate the 8 MVs by calculating their relative MSE and elect the candidate with minimal MSE. But sometimes the best MV still can not reconstruct the good enough quality as those from spatial concealment. Here, we set a threshold for switching both temporal and spatial concealment to provide the best quality. The threshold comes from experimental trials.

5. Hardware Implementation

Due to the limited computation power and resources, parts of encoder and decoder workload need the coprocessors to share for throughput enhancement. According to the profile in Fig. 7, motion estimation (ME) and inverse discrete cosine transform (IDCT) occupies the largest part in encoder and decoder respectively. We will use ARM platform with FPGA to show this architecture.

5.1. ME in encoder end

In general, motion estimation occupies 50%~80% in whole workload of encoder. To speedup this module will have significant improvement. To design such a co-processor, we choose ABME (All Binary Motion Estimation) [16] which is one kind of binary pyramid motion search to be our ME algorithm. Since this algorithm is just to keep only one bit for motion searching, the memory bandwidth and matching computation is expected to be lowered down significantly. In the original algorithm, some design tips will bring us difficulties in hardware realization, and the goal for modification is to make the data flow regular. The modified parts include fine tune, source change for motion vectors in level-2 stage, image padding shown in Fig. 8, and the results can be shown in Table 1.

Basically, the design flow can be separated into 4 parts including binarization, level-1 search, level-2 search, and level-3 search, and each level can follow Fig. 9. For easy hardware implementation and data reuse issues, MB (MacroBlock) based modules are design. The cost is shown in Table 2 with total gate of 68,494, and total running cycles for a MB is at least 283.

5.2. IDCT in decoder end

In decoder, IDCT and motion compensation (MC) is the largest 2 parts, but IDCT has much regular data flow. To implement IDCT in our decoder, a butterfly architecture shown in Fig. 10 is used. With 16 times of one-dimension IDCT, including row and column operation, a MB can be finished in 18 cycles. The total hardware cost is 48,379.

5.3. Test platform

The ARM platform shown in Fig. 11 contains RISC, logic module (FPGA), Static memory interface (SMI), ABMA bus and host interface. The designed co-processors is put in logic module and communicated through AHB (ARM high performance bus). ARM RISC runs 130 MHz and 10 MHz for logic module, 32bit 33 MHz for AHB. The experiment results are 42% throughput improvement in encoder and 8.1% in decoder.

To analyze this experiment data, it still is far from real-time requirement. To target this in the future, ARM instruction level optimization will be taken into consideration.

6. Experiment Results

To evaluate the performance of the proposed error resilience system, we give the following 3 experiments as an example. To target on the VOD application, we test 3 kind of bitrate including 256, 550, and 700 kilo-bps according to different network condition. For each experiment, we list 4 types of error resilience system tabulated in Table 3. Type 1 represents the original encoder and decoder system with default error resilience tools, including RM and zero-motion-vector spatial copy to the lost MB. Type 2 enables proposed intra refresh algorithm in the decoder. Type 3 enables proposed error robustness and error concealment algorithm. Type 4 is our proposed error system with intra-refresh in encoder and error concealment in decoder. To test 3 kind of different network condition, including packet loss rate (PLR) of 1%, 5% and 10% with uniform dropping probability model. Test sequence is encoded with one I-VOP's and 99 P-VOP's with Foreman sequence.

The test results are shown in Tables 4-6. We can find the Type 2 can get a gain over Type 1 with 3.64-8.45 dB in PSNR. Type 3 can achieve a gain over Type 1 with 5.16-10.62 dB in PSNR. Based on error-resilience decoder and rate-distortion intra refresh encoder, we can get a gain over Type 1 with 6.33-13.22 dB in PSNR. Fig. 12 shows the quality improvement over the other 3 types

and the objective quality is much better.

7. Conclusion

We have proposed a whole error resilience system on MPEG-4 video encoder and decoder. To lower the compact of error propagation, we applied intra refresh in the encoder. To achieve the best rate distortion balance, the Lagrange formula is modified to fit the real need. Modeling the network condition to fit the real decoder situation brings obvious improvement. In the decoder, the hybrid concealment algorithm improved the reconstructed image quality. With the perfect match of simulation in the encoder and decoder, it was proved to get at least 1 dB in PSNR more gain.

For the demo system on Linux platform, the future work is to optimize the elementary functions in speed and to integrate the real-time encoder and decoder into Linux-based platform the ARM-9 device.

Table 1. Experiment results with slight modification to ABME. (Search range of +/- 16)

Sequence (target bitrate, size)	Method	Y_PSNR(db)	Total Bits	Δ Y_PSNR(db)
Foreman (112kbps, CIF)	FSMB	29.48	1141680	
	ABME [16]	28.92	1174416	-0.56
	AMBE_ modified	28.96	1161840	-0.52
Akiyo (112kbps, CIF)	FSMB	40.83	1115456	
	ABME [16]	40.81	1115448	-0.02
	AMBE_ modified	40.83	1115384	-0.00
Mother-Daughter (112kbps, CIF)	FSMB	38.69	1112800	
	ABME [16]	38.31	1117048	-0.38
	AMBE_ modified	38.31	1116648	-0.38
Coastguard (112kbps, CIF)	FSMB	26.50	1120520	
	ABME [16]	26.33	1260528	-0.17
	AMBE_ modified	26.33	1242728	-0.17

Table 2. Hardware cost in each level design and binarization.

	Level-1	Level-2	Level-3	Binarization
Cycles per MB	17	11-13	12	228
PE no.	4	1	3	7
Gate count	5,163	14,717	22,435	26,179
Total gates	5,163 + 14,717 + 22,435 + 26,179 = 68,494			

Table 3. The 4 types of system with embedded ER tools.

Type	Encoder	Decoder
1	Resynchronization markers	Zero motion for P-VOP
1a	Resynchronization markers	Zero motion for P-VOP and spatial copy for I-VOP
2	Intra-refresh	Zero motion for P-VOP
2a	Resynchronization markers	Zero motion for P-VOP and spatial copy for I-VOP
3	Resynchronization markers	Proposed hybrid concealment
4	Intra-refresh	Proposed hybrid concealment

Table 4. The reconstructed image quality in PSNR(dB) for Foreman with 260 kbps.

PLR[2]		Type 1	Type 1a	Type 2	Type 2a	Type 3	Type 4
0%	PSNR[3]	33.59	-	-	-	-	-
	Gain[1]	-	-	-	-	-	-
1%	PSNR	25.58	28.26	30.40	31.14	30.74	31.91
	Gain[1]	0	2.68	4.82	5.56	5.16	6.33
5%	PSNR	16.71	21.8	23.58	25.1	25.95	27.48
	Gain[1]	0	5.09	6.87	8.39	9.24	10.77
10%	PSNR	14.17	19.03	20.73	22.92	22.97	26.35
	Gain[1]	0	4.86	6.56	8.75	8.8	12.18

[1]Gain: the difference compared to PSNR of type 1 (unit:dB)

[2]PLR : packet loss rate

[3]unit: dB

Table 5. The reconstructed image quality in PSNR (dB) for Foreman with 550 kbps.

PLR[2]		Type 1	Type 1a	Type 2	Type 2a	Type 3	Type 4
0%	PSNR[3]	36.94	-	-	-	-	-
	Gain[1]	-	-	-	-	-	-
1%	PSNR	25.61	28.87	32.44	33.36	31.97	34.92
	Gain[1]	0	3.26	6.83	7.75	6.36	9.31
5%	PSNR	16.83	22	25.28	27.09	26.65	29.78
	Gain[1]	0	5.17	8.45	10.26	9.82	12.95
10%	PSNR	14.15	19.64	21.79	24.25	23.90	27.37
	Gain[1]	0	5.49	7.14	10.1	9.75	13.22

[1]Gain: the difference compared to PSNR of type 1 (unit:dB)

[2]PLR : packet loss rate

[3]unit: dB

Table 6. The reconstructed image quality in PSNR (dB) for Foreman with 700 kbps.

PLR[2]		Type 1	Type 1a	Type 2	Type 2a	Type 3	Type 4
0%	PSNR[3]	38.04	-	-	-	-	-
	Gain[1]	-	-	-	-	-	-
1%	PSNR	26.76	30.1	30.40	33.47	33.83	34.90
	Gain[1]	0	3.34	3.64	6.71	7.07	8.14
5%	PSNR	16.51	21.6	23.58	28.05	26.29	31.31
	Gain[1]	0	5.09	7.07	11.54	9.78	14.80
10%	PSNR	13.92	19.41	20.73	25.35	24.54	28.48
	Gain[1]	0	5.49	6.81	11.43	10.62	14.56

[1]Gain: the difference compared to PSNR of type 1 (unit:dB)

[2]PLR : packet loss rate

[3]unit: dB

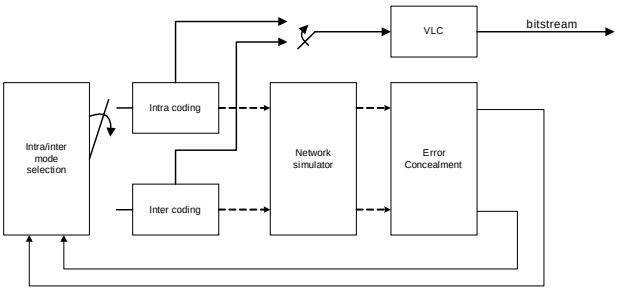


Fig. 1. The function blocks of MPEG-4 error resilience encoder.

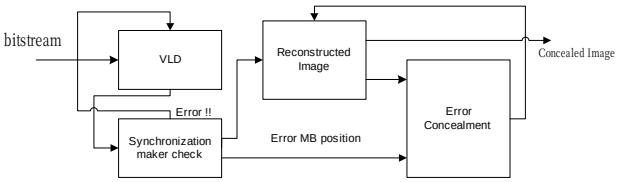


Fig. 2. The function blocks of MPEG-4 error resilience decoder.

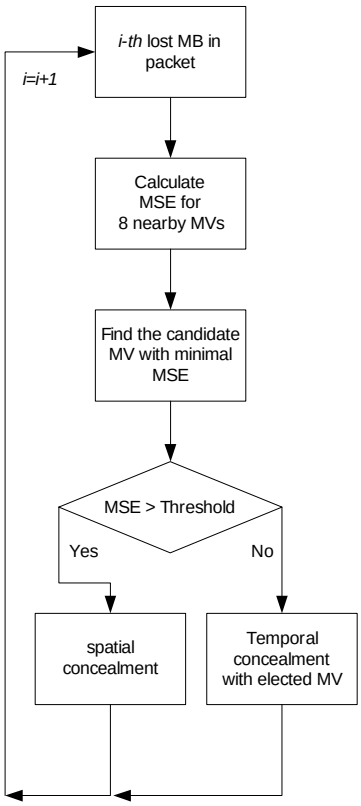


Fig. 4. The hybrid error concealment decoding flow.

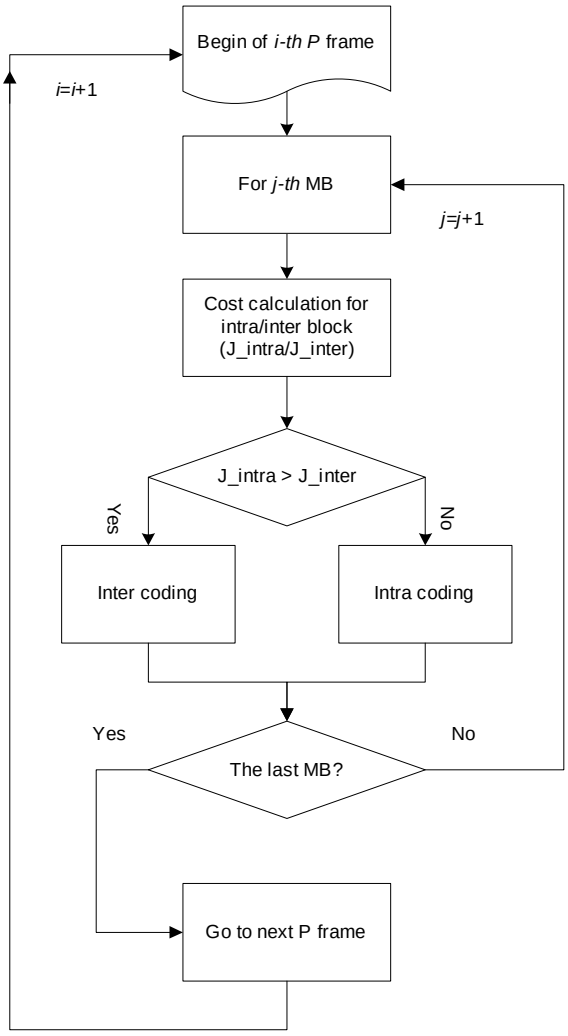


Fig. 3. The RDIR encoding flow

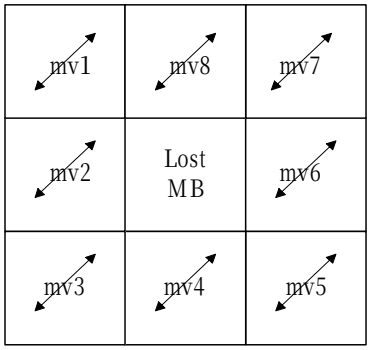


Fig. 5. Temporal concealment using neighboring motion vectors

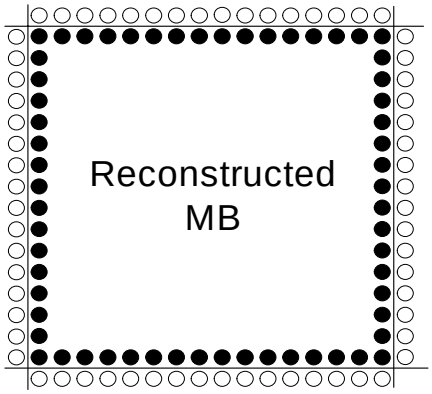
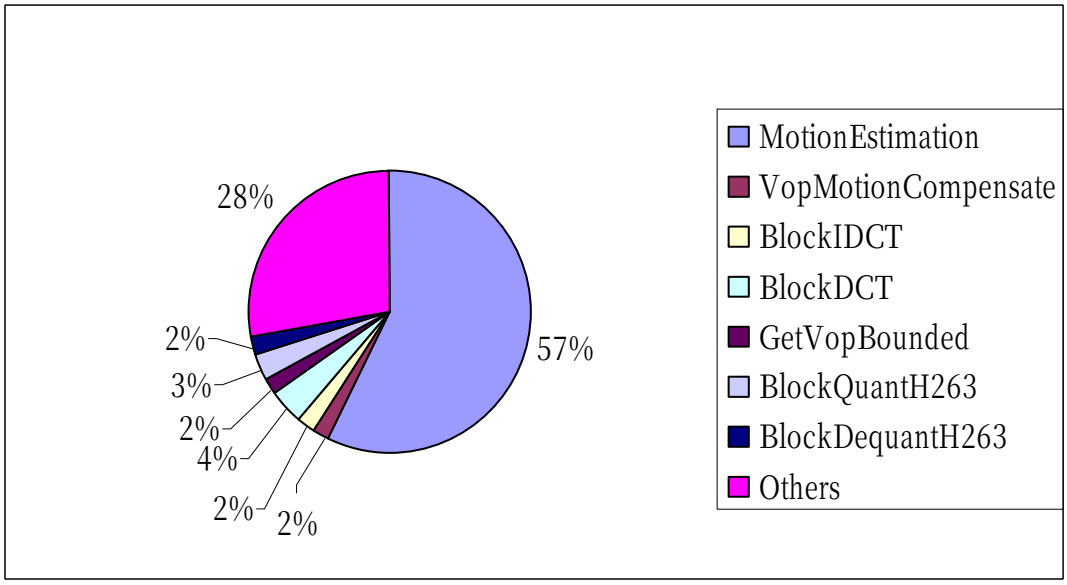
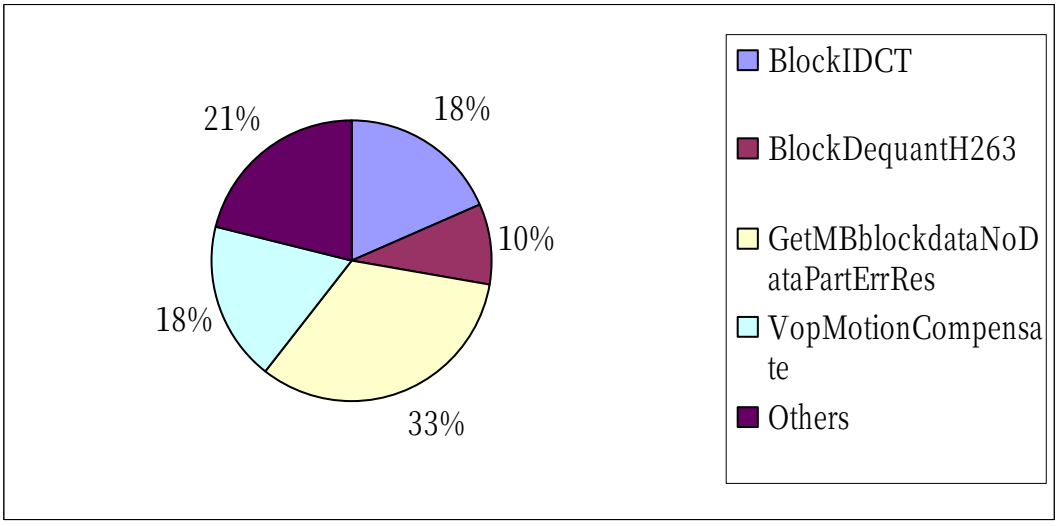


Fig. 6. The pixels used for calculating MSE



(a) Encoder



(b) Decoder

Fig. 7. The profiling of proposed MPEG-4 error resilience system. (a) Encoder (b) Decoder.

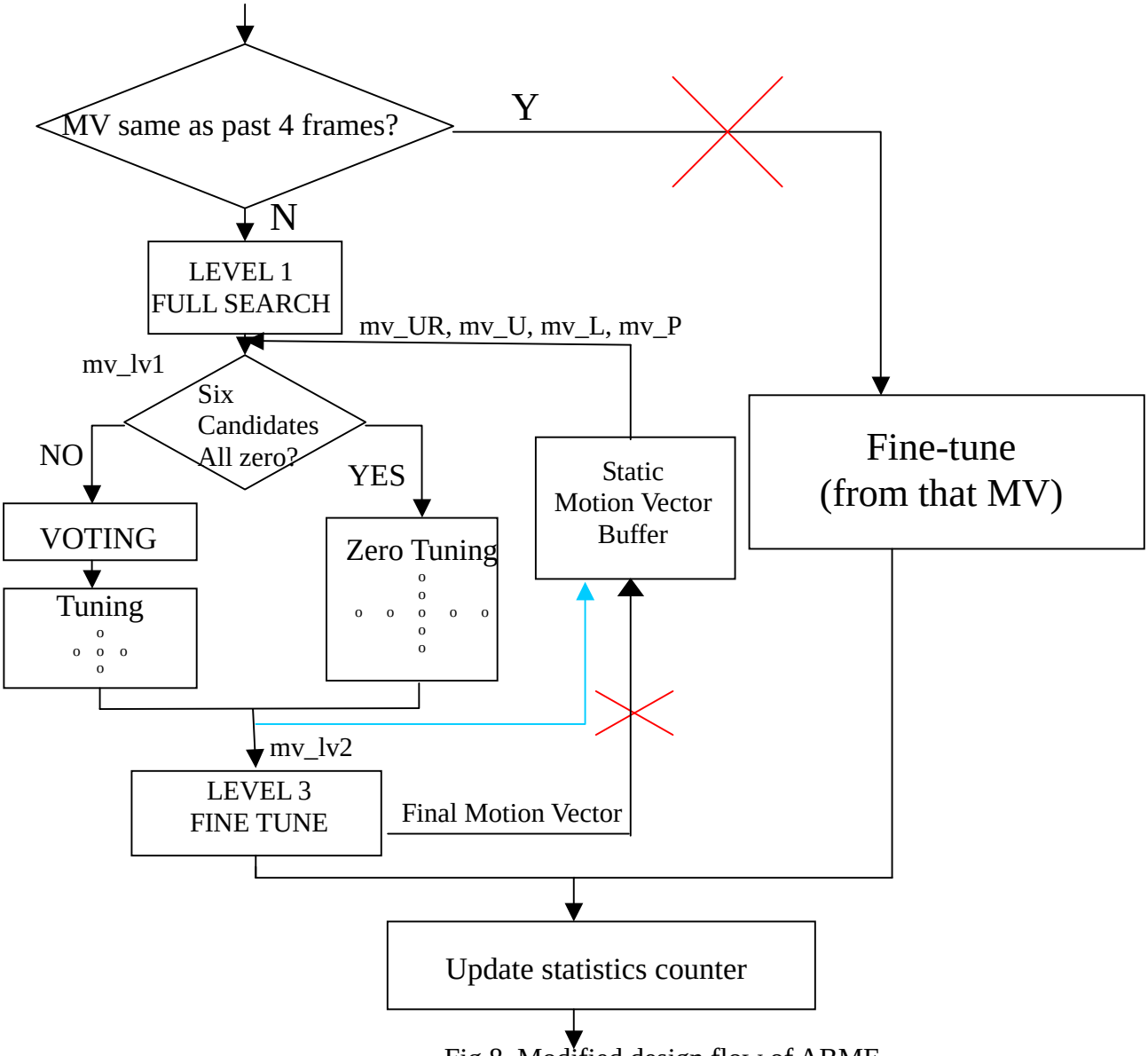


Fig 8. Modified design flow of ABME.

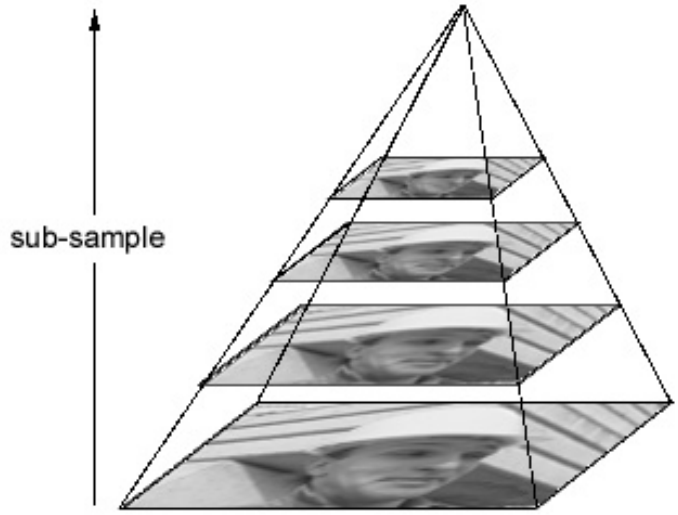


Fig 9. The structure of binary pyramid search.

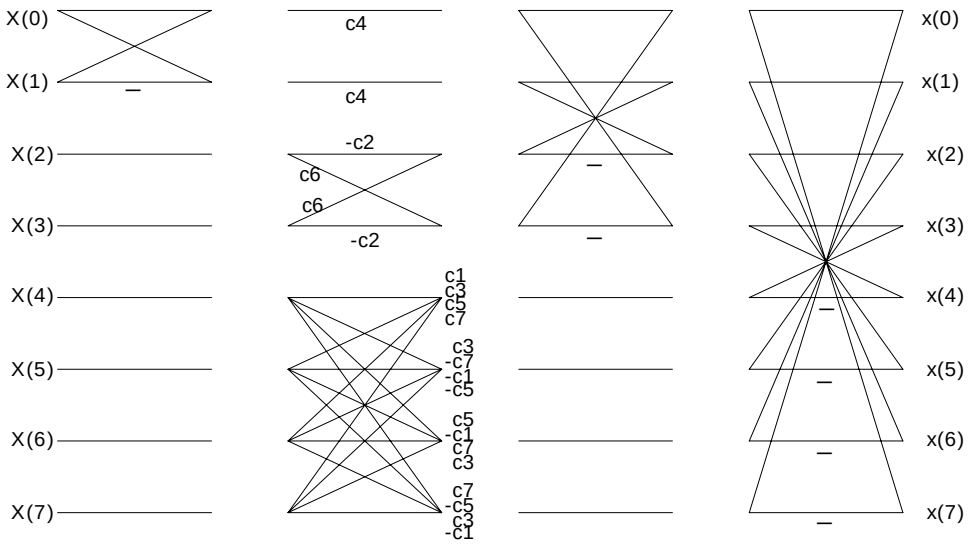


Fig 10. The flow diagram of IDCT.

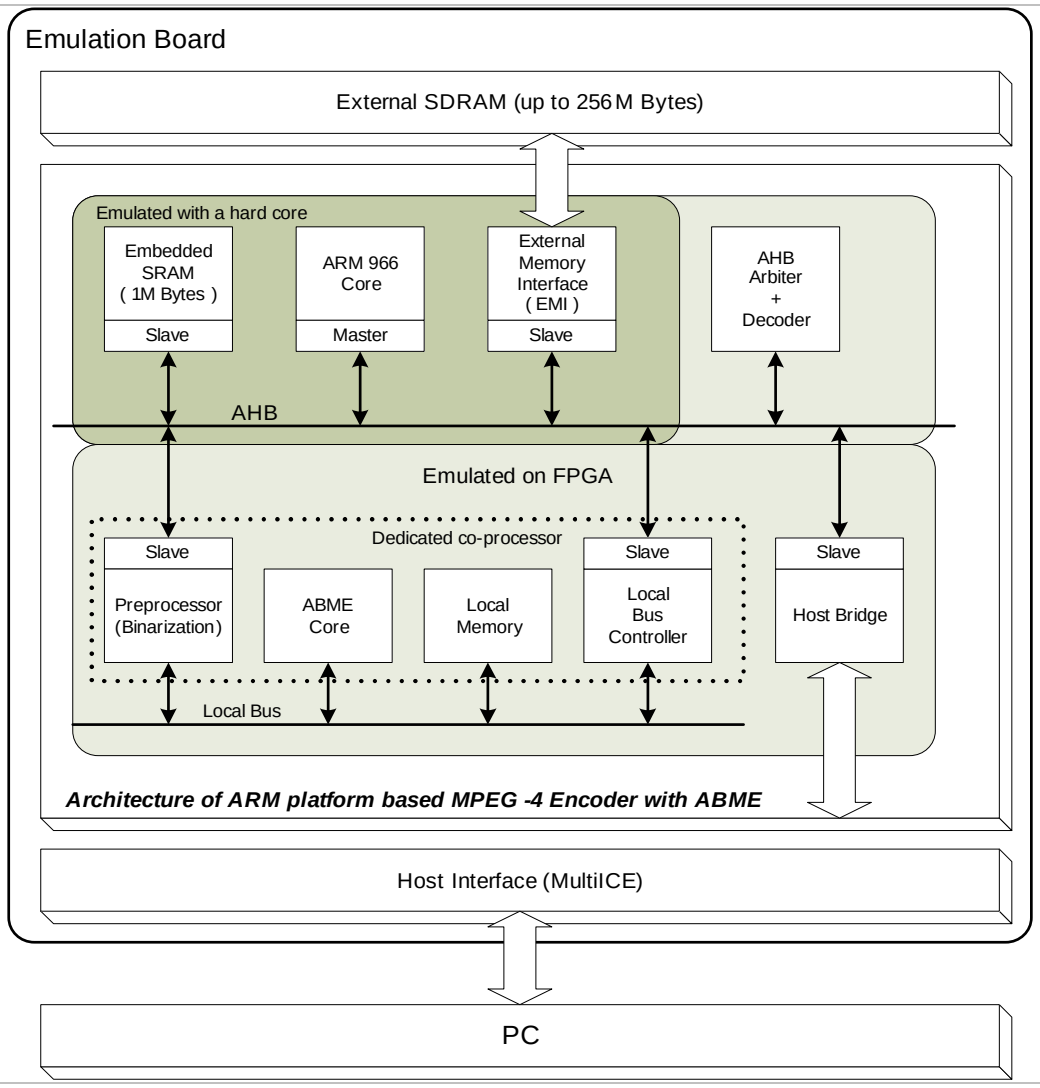
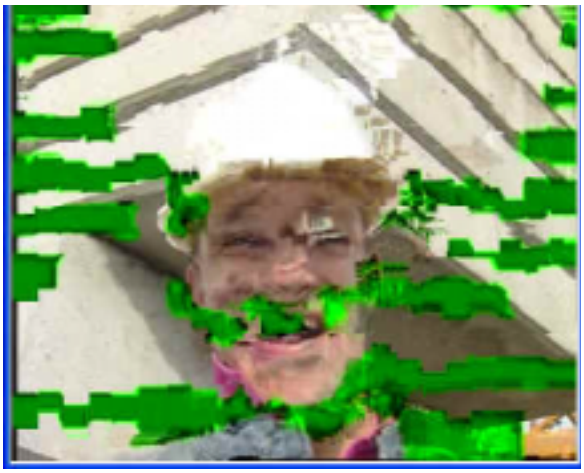


Fig. 11. The overview of the proposed ARM platform.



(a)

(b)

(c)

(d)

(e)

(f)

Fig 12. The image quality comparison for the 4 types of error resilience system. (a) Type 1 (b) Type 1a (c) Type 2 (d) Type 2a (e) Type 3 (f) Type 4. (No. 10th frame, PLR 10%, bitrate 700kbps)

8. References

[1] A. Dagiuklas and M.Ghanbari, “Packet video transmission in an ATM network using forced frame refreshment,” *Proc. Third IEEE International Conference on Electronics, Circuits, and Systems*, vol. 2, pp.784 -787, Oct. 1996.

[2] L. Favalli, C. Fraschini and A. Mecocci, “A low refresh-rate video sequences compression technique using quadtrees and adaptive spatial sampling,” *Proc. Sixth International Conference on Image Processing and Its Applications*, vol. 1, pp.41 -45, July 1997.

[3] E. Steinbach, N. Farber and B. Girod, ”Standard compatible extension of H.263 for robust video transmission in mobile environments,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, pp. 872 -881, Dec. 1997.

[4] J. Y. Liao and J. Vilasenor, “Adaptive intra block update for robust transmission of H.263,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, pp. 30 -35, Feb. 2000

[5] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, “Error resilient video coding techniques,” *IEEE Signal Processing Magazine*, vol. 17, pp. 61-82, July 2000.

[6] R. Talluri, “Error resilient video coding in the ISO MPEG-4 standard,” *IEEE Communications Magazine*, vol. 36, pp. 112 -119, June 1998.

[7] D.-S. Luis and P. Fernando, “Error resilience and concealment performance for MPEG-4 frame-based video coding,” *Signal Processing: Image Communication*, 1999

[8] Y.-L. Lee, C.-N. Wang and Tihao Chiang, “An MPEG-4 error resilient decoder,“ *Proc. of Workshop on Consumer Electronics*, 2001.

[9] S.-H. Wang, C.-N. Wang, Tihao Chiang and H. Sun, “AHG report on editorial convergence of MPEG-4 reference software,” ISO/IEC JTC1/SC29/WG11, M8884, 2002.

[10] Wen-Jeng Chu and Jin-Jang Leou, “Detection and concealment of transmission errors in H.261 images,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 8, pp. 74 -84, Feb. 1998

[11] S. Valente, C. Dufour, F. Groliere and D. Snook, “An efficient error concealment implementation for MPEG-4 video streams,” *IEEE Transactions on Consumer Electronics*, vol. 47, pp. 568 -578, Aug. 2001.

[12] R. Zhang, S.L. Regunathan and K. Rose, “Video coding with optimal inter/intra-mode switching for packet loss resilience,” *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 966 –976, June 2000.

[13] G. Cote, S. Shirani and F. Kossentini, “Optimized mode selection and synchronization for robust video communications over error-prone networks,” *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 952-965, No. 6, June 2000.

[14] S. Kaiser and K. Fazel, “Comparison of error concealment techniques for an MPEG-2 video decoder in terrestrial TV-broadcasting,” *Signal Processing: Image Communication* 14, pp. 655~676, 1999.

[15]S. Aign and K. Fazel, “Error detection &concealment measures in MPEG-2 video decoder,” *Proc. International. Workshop on HDTV'94*, Torino, Italy, October 1994.
[16] Jeng-Hung Luo, Chung-Neng Wang, Tihao Chiang, “A novel all-binary motion estimation (ABME) with optimized hardware architectures,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, Aug. 2002

三、　計畫成果自評

本計畫從事基於 MPEG-4 視訊壓縮標準之壓縮與解壓縮器在抗噪性與錯誤回復力方面和即時性之研究。除此，為了改善影像品質，我們也對回復的影像給予簡單的錯誤補償。為了完成在 ARM 平台上即時性的壓縮與解壓縮器，我們也採用一些快速的演算法與高效率的記憶體存取方法來改善壓縮與解壓縮器執行效能，並與予實現計算複雜度高的硬體。在國科會的支持下，我們較產業早一步對這些規格標準加以探討，發展其中關鍵技術。如第四部份研發成果所示，本計畫已獲得相當豐碩成果，發表學術論文與專利申請，與原訂目標相符。

協同參與本項計畫的杭學鳴、王聖智老師，並在業界廠商合作計畫補助旅費情況下參與 MPEG 標準會議，並於每次會議兩週後即舉行公開之說明會對業界提供最新之訊息。現已提建議案多項，例如目前我們在 MPEG 標準會議進行之主要工作項目有 MPEG-4 Part 7 Optimised Reference Software 以及 MPEG-21 Part 12 Multimedia Test Bed for Resource Delivery。（我們參與 MPEG 標準會議的技術開發與活動經費，亦受到交通大學李立台揚網路研究中心與多媒體標準資源共享等計畫之贊助。）

此外，更直接並且對國內產學界更直接且有價值的貢獻將是人才訓練。同學們在學校階段已熟悉較前瞻的世界多媒體標準 MPEG-4、硬體設計、軟硬體整合最佳化、抗噪性與錯誤回復力、與系統整合與最佳化，畢業後進入產業， 直接投入產業界開發新產品，以提昇我國多媒體技術之研發與整合能力。

綜合評估：本計畫產出相當多具有學術與應用價值的成果，並積極參與國際 MPEG 標準會議，將我國人研發成果推廣到國際舞台。此外亦達到媒體技術之研發與整合之人才培育之效，整體成效良好。

四、 研發成果

1. Journal papers (1)

   a、 J. H. Luo, C. N. Wang, and Tihao Chiang, "A novel all binary motion estimation with optimized hardware architectures, " *IEEE Transaction on Circuits and Systems for Video Technology—Special Issue on Multimedia Implementation, Aug.* 2002.

2. International conference papers (3)

   a、 J. H. Luo, C. N. Wang, and Tihao Chiang, "A Novel All Binary Motion Estimation (ABME), "*in Proceeding of ISCAS02*, 2002.

   b、 Shih-Hao Wang, Wen-Hsiao Peng, Yuwen He, Guan-Yi Lin, Chen-Yi Lin, Shih-Chien Chang, Chung-Neng Wang, and Tihao Chiang, "A Platform-Based MPEG-4 Advanced Video Coding (AVC) Decoder with Block Level Pipelining", Accepted by PCM2003.

   c、 Chih-Hung Li, Chung-Neng Wang, and Tihao Chiang, "A VBR Rate Control Using MINMAX Criterion for Video Streaming, " *in Proceeding of PCM02 (HsinChu, Taiwan)*, Dec. 2002, pp.831-838.

3. Domestic conference papers (4)

   a、 Yue-Lin Lee, C.-N. Wang, and Tihao Chiang, "An MPEG-4 Error Resilient Decoder, " *in Proceeding of WCE 2001*.

   b、 S.-H. Wang, C.-N. Wang, and Tihao Chiang, "An optimized MPEG-4 Reference Software," *in Proceeding of Lee Center's Workshop*, 2001.

   c、 S.-H. Wang, G.-Y. Lin, C.-N. Wang, X.-Y. Wu, and Tihao Chiang, "A Fast Variable Length Decoder for MPEG-4 Using Hierarchical Table Lookup," *in Proceeding of WCE 2002*.

   d、 Ming-Yen Huang, Tzu-Liang Su, Shih-Hao Wang, Chung-Neng Wang, and Tihao Chiang, "An Error Resilient System for Streaming MPEG-4 Video over the Internet," *Accepted by WCE2003, Oct. 2003*.

4. Patents (3)

   - ABME (細節請參閱可供推廣之研發成果資料表(一))

   a、 C.-H. Luo, G.-M. Lee, C.-N. Wang, and Tihao Chiang, "Method And Apparatus For Estimation With Binary Representation," Filed in U.S patent (IAM10/301415), February 26th, 2003.

   b、 C.-H. Luo, G.-M. Lee, C.-N. Wang, and Tihao Chiang, "Method And Apparatus For Estimation With Binary Representation," Filed in Japan patent (2003-048958), November 11th, 2002.

   c、 C.-H. Luo, G.-M. Lee, C.-N. Wang, and Tihao Chiang, "在視訊影像編碼中作為動態預測的裝置與方法," Filed in R.O.C patent(92104361), September 27, 2003

5. Master thesis (2)

   a、 C.-H. Luo and Tihao Chiang, “On a Fast Motion Estimation Algorithm, ” Master Thesis, Dept. of Electronic Engineering, National Chiao Tung University, June, 2001

   b、 Wei-Lun Tao and Tihao Chiang, **“**ARM 平台的全二元動量估測架構之設計 An All-Binary Motion Estimation Architecture Design on ARM-Based Platform”, Master Thesis, Department of Electronics Engineering, National Chiao Tung University, Oct. 2003.

6. MPEG contributions (19)

   a、 C.-N. Wang, J.-H. Luo, and Tihao Chiang, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* N3674: Description of Rate Control Core Experiments Q6, ” Oct. 2000 (La Baule).

   b、 J.-H. Luo, C.-N. Wang, and Tihao Chiang, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M6889: Results of Rate Control Core Experiments Q6, ” Jan. 2001 (Pisa)

   c、 C.-N. Wang, C.-Y. Lee, Tihao Chiang, and H.-M. Hang, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M7784: Work Plan on Software Integration and Verification for MPEG-4 14496-5 Reference Software, ” Oct. 2001 (Pattaya, Thailand).

   d、 S.-H. Wang, C.-N. Wang, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8041: AHG report on editorial convergence of MPEG-4 reference software, ” Mar. 2002 (Jeju Island, Korea).

   e、 C.-N. Wang and Tihao Chiang, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8407: Draft Call for Proposals on Fast Motion Estimation for 14496-10 JVT, ” May 2002 (Fairfax, VA, USA).

   f、 S.-H. Wang, Yao-Chung Lin, C.-N. Wang, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8408: AHG report on editorial convergence of MPEG-4 reference software, ” May 2002 (Fairfax, VA, USA)

   g、 S.-H. Wang, C.-N. Wang, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8603: AHG report on editorial convergence of MPEG-4 reference software, ” July 2002.

   h、 S.-H. Wang, C.-N. Wang, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8884: AHG report on editorial convergence of MPEG-4 reference software, ” Oct. 2002.

   i、 C.-N. Wang, Tihao Chiang, and Jens-Rainer Ohm, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8885: MPEG-4 Visual: Updated List of Problems Reported, ” Oct. 2002.

j、 S.-H. Wang, C.-N. Wang, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M8886: Proposed Text of Proposed Draft Technical Reports of ISO/IEC PDTR 14496-7 for Optimized Simple Profile Reference Software, ” Oct. 2002.

k、 S.-H. Wang, C.-N. Wang, G.-Y. Lin, Tihao Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M9073: AHG report on editorial convergence of MPEG-4 reference software, ” Dec. 2002

l、 C.-N. Wang, Y.-S. Tung, Tihao Chiang, and Jens-Rainer Ohm, “ISO/IEC JTC1/SC 29/WG 11 *14496-2* M9181: MPEG-4 Visual: Updated List of Problems Reported, ” Dec. 2002.

m、 S.-H. Wang, C.-N. Wang, Yi-Shin Tung, T. Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 M9632: AHG report on editorial convergence of MPEG-4 reference software, ”July 2003.

n、 C.-N. Wang, Y.-S. Tung, T. Chiang, and Jens-Rainer Ohm, “ISO/IEC JTC1/SC 29/WG 11 M9484: MPEG-4 Visual: Updated List of Problems Reported, ” Mar. 2003.

o、 S.-H. Wang, C.-N. Wang, Yi-Shin Tung, T. Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 M9355: AHG report on editorial convergence of MPEG-4 reference software, ” Mar. 2003

p、 C.-N. Wang, Y.-S. Tung, T. Chiang, and Jens-Rainer Ohm, “ISO/IEC JTC1/SC 29/WG 11 M9763: MPEG-4 Visual: Updated List of Problems Reported, ” July 2003.

q、 C.-N. Wang and T. Chiang, “ISO/IEC JTC1/SC 29/WG 11 M9764: Improvement of Optimized MPEG-4 14496-2 Simple Codec with EPFL SIT Analyzer, ” July 2003.

r、 S.-H. Wang, C.-N. Wang, Yi-Shin Tung, T. Chiang, and H.F. Sun, “ISO/IEC JTC1/SC 29/WG 11 M9951: AHG report on editorial convergence of MPEG-4 reference software, ” October 2003.

s、 C.-N. Wang, Y.-S. Tung, T. Chiang, and Jens-Rainer Ohm, “ISO/IEC JTC1/SC 29/WG 11 M10173: MPEG-4 Visual: Updated List of Problems Reported, ” October 2003.

7. JVT documents (1)

a、 G.-M. Lee, C.-N. Wang, and Tihao Chiang, “JVT-F079: Cross check results for JVT-F017, “ Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6), Dec. 2002.

附件二

可供推廣之研發成果資料表(一)

☐ 可申請專利　☑ 可技術移轉　日期：92 年 10 月 30 日

國科會補助計畫	計畫名稱：單晶片無線多媒體資訊家電之設計與製作(3/3) -- 子計劃四 ： 單晶片無線多媒體通訊系統 計畫主持人：蔣迪豪 交通大學電子工程系所 副教授 計畫編號：NSC－91-2218-E-009-005 學門領域：SoC
技術/創作名稱	All Binary Motion Estimation (ABME)
發明人/創作人	羅正弘，李鑑明，王俊能，蔣迪豪
技術說明	中文： 英文： 1. We present a fast motion estimation algorithm using only binary representation, which is desirable for both embedded system software optimization and hardware implementation with parallel architectures. 2. Additionally, our fast motion estimation algorithm employs the other two alternative Boolean operations instead of SoD (using XOR operation) as interblock similarity measures. The new measure is Sum of One (SoO) and Sum of Zero (SoZ). The SoO uses the AND operation for similarity checking and the SoZ takes the NOR operation for similarity checking. 3. Finally, our fast motion estimation algorithm accomplishes the exhaustive search with the sequentially arranged binary data in each pyramidal level, which provides a feasible hardware implementation. The experimental results show that the proposed algorithm is applicable for smaller picture size at low bitrates such as MPEG-4 and H.263 applications. It is also useful for the applications of larger picture size at high bitrates such as MPEG-2 applications
可利用之產業及可開發之產品	1. Films using MPEG or H.26X, Global motion estimation video encoding techniques 2. Surveillance 3. Multimedia like DVD, VCD, HDTV. 4. Content providers 5. MPEG video related software encoder 6. Video indexing using motion features 7. Wireless communication-based appliances

技術特點	This invention has been made to overcome the above-mentioned drawbacks of conventional motion estimation. The primary object is to provide a method for motion estimation with all binary representation for video coding. Accordingly, a binary pyramid having three binary layers of video images is constructed. The first binary layer is first searched with a criterion based on bit-wise sum of difference to find a first level motion vector. Six motion vector candidates are used to determine a motion vector in the second binary layer. Finally, a search in the third binary layer according to the second layer motion vector generates a final motion vector. In the present invention, the construction of the binary pyramid includes filtering, binarization and decimation. The precise edge information is extracted based on the spatial variation within a small local area of an image to provide all binary edge information without having to use any integer layer. In the first level search, the search is performed within a ± 3 pixel refinement window. In the second level search, based on the spatio-temporal dependencies that exist among blocks, ABME calculates the ranges of two dimensional 8×8 motion offsets $([R^x_{\min}, R^x_{\max}],[R^y_{\min}, R^y_{\max}])$ through the six motion vector candidates from the current and previous frames. The refinement window in the second level has thus covered the dominant ranges of the search area with dimension $(R^x_{\max}+R^x_{\min})\times(R^y_{\max}+R^y_{\min})$ around the mean vector of the six motion vectors. We then perform the full-search XOR Boolean block matching with $(R^x_{\max}+R^x_{\min})\times(R^y_{\max}+R^y_{\min})$ pixels for refinement at the second level. Similarly, the resultant motion vector candidate will be passed onto the next binary level. In the third level, the search is performed within a ± 2 pixel refinement window. At each level, the search and determination of the best motion vector is based on a criterion of minimum bit-wise sum of difference using XOR block matching. It is also an object of the invention to provide an apparatus for motion estimation for video encoding. Accordingly, the apparatus comprises a binary pyramid construction module, a first level search module, a second level search module, and a third level search module. Each level search module includes a data loading module, a bit alignment module, and an XOR block matching module. The binary pyramid construction structure further comprises a filtering module, a binarization module and a decimation module. Each XOR block matching module further includes a table lookup sub-module and a bit-wise sum of difference (SoD) sub-module.

	The motion estimation of this invention is feasible for pipelined architectures. The method of motion estimation can be implemented in various architectures including general-purpose architectures such as x86, single instruction multiple data (SIMD) architectures using Intel's MMX technology, and systolic arrays. The pipelined architecture of the invention contains three major common modules including the integrated construction, compact storage, and parallel block matching. The invention uses a MPEG-4 reference video encoder and employs a macroblock with size 16×16 for block matching to show the performances. According to the experimental results, it not only has the benefits of low computational complexity and low memory bandwidth consumption but also is insensitive to search range increase. System designer can choose better binarization methods to further improve the visual quality. In addition, various optimization methods can be developed for specific platforms with different register size. The invention thus is more flexible than other motion estimation method. From the operation counts, the motion estimation of this invention is very desirable for software implementation on a general-purpose processor system. It can be realized by a parallel-pipelined implementation for ASIC design and allows tradeoffs between Silicon area, power consumption and visual quality during the hardware design phase.
推廣及運用的價值	

※ 1.每項研發成果請填寫一式二份，一份隨成果報告送繳本會，一份送 貴單位研發成果推廣單位（如技術移轉中心）。

※ **2.本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。**

※ 3.本表若不敷使用，請自行影印使用。