

台灣高畫質電視 HDTV 頻道

H.264 影像壓縮格式

主客觀評量測試評估報告

公共電視台 工程部

2010/04

目 錄

壹、 緒論	6
1.1. 研究動機與目的	6
1.2. 章節概要	6
貳、 台灣 DVB-T 單頻網路傳輸率及高畫質電視通道說明	8
2.1. 台灣無線數位電視網路現況說明	8
2.2. 以固定接收做為高畫質電視網路條件下的規劃	9
2.3. DVB-T 通道實現理想中的中高資料流量的固定接收形式	10
2.4. 小結	11
參、 影像主/客觀評量測試方法	12
3.1. 說明	12
3.2. 影像品質評估方法	12
3.3. 影像品質的主觀評量	12
3.4. 影像品質客觀評量說明	14
3.5. 主觀評量實施依循規範	16
3.6. 主觀評量系統說明	19
3.7. 評分方法和評分標準	23
3.8. 主觀評量的數據統計方法	24
肆、 主觀評量作業程序說明及測試結果分析	27
4.1. 測試系統規劃：	27
4.2. 主觀影像測試方法	27
4.3. 測試結果統計分析：	33
伍、 影像品質客觀評量標準與測試結果分析	38
5.1. 客觀評量	38
5.2. 信號源頭端影像品質的客觀評量	39
5.3. 影像品質的 PSNR 測量	40
5.4. 測試片段 PSNR 量測值及分析	40

5.5. 小結：	43
陸、 結論	45
6.1. 對我 HiHD 頻道壓縮傳輸資料量的建議	45
6.2. 提高節目壓縮資料量，影像品質真能提高多少？	46
6.3. 國外相關研究分析對照及文獻說明	46
6.4. 未來技術演進，效率與品質會更好	48
柒、 參考文獻.....	51
附錄 LG LCD TV 42LG60 規格.....	54

表目錄

表格 1、英國 SkyHD 在 H.264 廣播用影像壓縮率提出的建議值.....	9
表格 2、不同調變參數所提供的資料傳輸量	9
表格 3、室外固定接收最低電場強度需求 (單位：dBuV/m).....	9
表格 4、種調變模式下影像有效位元速率	10
表格 5、全省 10 大站場強模擬計算面積及人口值.....	10
表格 6、兩種評量的影像 5 級評分表	14
表格 7、數位電視影像品質主觀評量觀看條件要求.....	17
表格 8、測試素材的選擇	20
表格 9、受測人員背景資料	28
表格 10、客觀 PSNR 值與主觀評量 MOS 對應表	43
表格 11、HiHD 頻道頻寬建議參考值	45
表格 12、H.264 的分級特性	49
表格 13、係數.....	49

圖目錄

圖表 1、我國數位電視頻譜分配示意圖	11
圖表 2、系統的原理方塊圖	19
圖表 3、測試階段的流程	22
圖表 4、評分表.....	23
圖表 5、影音片段顯示評分的順序.....	24
圖表 6、測試系統方塊圖	27
圖表 7、測試影像畫面的循環順序及時間長度.....	28
圖表 8、受測環境示意圖	29
圖表 9、原始畫面品質的評量統計圖	33
圖表 10、受測畫面品質評量統計圖	33
圖表 11、評量測試結果 DMOS.....	34
圖表 12、各測試片段的標準誤差值	35
圖表 13、受測片段三種不同壓縮率畫面比較(1).....	36
圖表 14、受測片段三種不同壓縮率畫面比較(2).....	37
圖表 15、基於解碼影像與基準影像差值的品質模型	38
圖表 16、信號源端的影像品質相對評估	39
圖表 17、Omnitek PQA 內部信號處理流程示意圖.....	41
圖表 18、客觀評量測試系統示意圖	41
圖表 19、測試片段在不同壓縮率的 PSNR(Y)值.....	42
圖表 20、測試片段在不同壓縮率的 PSNR(C)值.....	42
圖表 21、PSNR 量測結果跟 MOS 對應表比對圖	43
圖表 22、歐洲 EBU 提出的最低資料量需求	47
圖表 23、日本總務省網站在 HDTV 傳輸資料量中的建議值資料	48
圖表 24、ATEME 公司所提出壓縮器內 FPGA 硬體效能	48
圖表 25、製造商希望下一代 H.264 壓縮器的產品效能目標	49
圖表 26、H.264 壓縮技術應用於整個製播環境示意圖	50

壹、緒論

1.1. 研究動機與目的

2008年二月公共電視台利用台北萬里轉播站傳送「高畫質電視測試信號」，實驗為台灣地區高畫質電視首次以MPEG4/H.264格式壓縮節目信號，再經DVB-T無線電波進行發射與接收測試，同年5月公視開始於北高兩地開始進行大台北及高高屏地區高畫質電視「HiHD」頻道試播，我台灣無線數位電視開始進入高畫質電視播映的時代。

由於鄰近的日本早在20年以前就經由衛星直播方式，開始進行高畫質電視的服務，不論是由早期類比的「MUSE」系統或是現在以MPEG2格式的數位壓縮傳輸方式，我國很多觀眾對於日本的BS-Hi頻道等高畫質頻道並不陌生，而對於NHK以MPEG2近25M bps的資料量所表現的影像畫質感到非常滿意。

HiHD高畫質電視測試頻道，由我NCC核配一個頻道Ch.30(中心頻率:569Mhz)6Mhz頻寬來進行全省單頻網(Single Frequency Network)規劃建置，所以在網路上整個資料承載量受到部分條件上的限制，有部分的收視戶對於壓縮後的HiHD頻道上所表現的高畫質影像畫質不甚滿意進而存疑，認為是傳輸的資料量過低是造成畫面品質不良的主因。

由於目前電視廣播所採用的動態影像壓縮技術，不論是MPEG2或是MPEG4/H.264格式都是利用人眼的視覺弱點特性進行破壞性的運算而將原有的傳輸資料量大幅降低，在對人眼視覺系統(HVS)的研究中，專家們發現人眼對不同亮度/色彩信號的敏感程度不同。在背景亮度較高的區域，人眼對於失真的敏感程度則會大大降低。因此，編碼器可以利用人眼視覺的這一特性進行技術改進。每個人在觀看解壓縮後的畫面時感受也因人而異，也就是感受是一個主觀知覺的認定。

目前在全球數位電視網路建置中，自2008年起，高畫質電視的壓縮格式均採用新一代壓縮技術H.264格式，如法國、英國、巴西、香港、挪威、等國家陸續採用，相信日後也將廣泛的應用在各種影音內容的服務。本研究要探討的是高畫質電視影像在MPEG4/H.264壓縮傳輸的資料量多少才能被觀眾所接受?我們利用相關規範進行主觀及客觀兩方面的測試分析，以取得一個具有參考價值並做為現有高畫質電視傳輸的依據。

1.2. 章節概要

本研究第壹章節為前言，說明本研究的背景、動機及目的。

第貳章節是介紹及簡述目前台灣DVB-T單頻網路架構、傳輸率及高畫質電視通道等相關條件說明與敘述。

第參章節說明高畫質影像主/客觀評量方法，及如何以現有的相關技術標準為基準進行研究。

第肆章節進行主觀評量結果分析，依據ITU-R BT.500及BT.710等相關標準進行人員主觀評量的結果分析及驗證，並參考客觀評量的方式進行交叉比對其結果。

第伍章節為進行影像品質客觀評估標準與測試結果分析，以公視現有 Omni-Tek PQA 分析儀進行受測片段的 PSNR 量測，取得相關數據做為影像品質客觀分析使用，提供參考及相關討論。

第陸章節為本研究¹的總結，提出高畫質影像在 H.264 壓縮時建議的最佳傳輸速率提供參考及相關技術發展等討論。

¹本研究報告版權歸屬財團法人公共電視文化事業基金會所有，本會保有修改本報告內容的權利，不另行公告。

貳、台灣 DVB-T 單頻網路傳輸率及高畫質電視通道說明

自 1996 年秋的 DVB-T 數位電視標準推出後，於 1998 年在英國率先進入市場開始商業運作。歐規 DVB-T 與美規 ATSC 這兩項技術標準的重要區別是：DVB-T 具有多種可能的技術組合，可以“因地制宜”靈活適應各國各地區 DTTB(Digital Terrestrial Television Broadcasting)的實際需求；而 ATSC 則是單一的技術組合。DVB-T 初期著重在 SDTV 內容服務，而 ATSC 則著重 HDTV 一步到位的服務。

然而這些年來，隨著各國建置無線數位電視網路建置一一完成的時候，相關技術也越來越成熟與穩定，壓縮技術也日新月異，除了在 SDTV 的服務外，人們對於高畫質電視 HDTV 的服務需求也越來越多。所以如何在現有的網路上，發展更多更好的服務內容將會是技術成熟之後要更進一步的目標。

2.1. 台灣無線數位電視網路現況說明

DVB-T 標準的頻道頻寬有 6、7、8 MHz 共 3 種選擇項。由於台灣地區於 2002 年選擇 DVB-T 技術做為我國地面波數位電視標準，而我國在電視頻譜的劃分上因早期受美規電視頻道的劃分所影響，所以我國在 DVB-T 的頻寬也選用了 6Mhz 頻道頻寬，但是卻造成日後我國在發展數位電視上諸多的困擾如下：

- DVB-T 發展初期，全世界僅台灣採用 6Mhz 頻寬的數位電視網路建置，造成國外許多發射機、測試儀表及機上盒等等製造供應廠商初期無法供應 6Mhz 頻寬設備之窘境，直至這兩年來情況才較有改善。
- 同樣調變參數的條件下，6Mhz 頻寬比 8Mhz 頻寬所能提供的資料承載率 (Payload) 要少 1.2M~8M bps，抗都卜勒效應(Anti-Doppler Effect)的能力也較差，行動接收的能力比 8Mhz 低，所以使得我國在使用同樣 DVB-T 技術相對資源變少。

而目前五家無線電視公司，其原有全台類比電視廣播站台的電波覆蓋範圍是多年前就規劃形成的(主要的服務範圍考量當地地理因素以及人口集中的區域為主)，所以在台灣地區地面波數位電視網路建置時，大部分也是以原站台做為規劃主軸，最大差異之一是以單一頻率(Signal Frequency Network)做為全島的網路規劃，也因此數位電視發射機的調變模式受到很大的限制，而使得整體數位網路的資料傳輸量也受到影響，以目前 5 家無線電視台的地面波數位電視台的建置來看，調變的參數都是 16QAM，CR:2/3，GI:1/4，在這個條件下整體的 Payload 為 9.9M bps。

電視台	發射機調變參數	承載量 (Payload)	服務內容
TTV	16QAM,GI:1/4,CR:2/3	9.9M bps	SDTV x3 (每個頻道資料量約 3M bps)
CTV	16QAM,GI:1/4,CR:2/3	9.9M bps	SDTV x3 (每個頻道資料量約 3M bps)
CTS	16QAM,GI:1/4,CR:2/3	9.9M bps	SDTV x3 (每個頻道資料量約 3M bps)

PTS	16QAM,GI:1/4,CR:2/3	9.9M bps	SDTV x3 (每個頻道資料量約 3M bps)
FTV	16QAM,GI:1/4,CR:2/3	9.9M bps	SDTV x3 (每個頻道資料量約 3M bps)

如果要用數位電視網路來傳送 HDTV 的節目內容時，若以 MPEG2 格式來進行信號即時壓縮，大約需要 20~25Mbps 的傳輸資料量才足夠傳送一個頻道的 HDTV；若依上述的數位電視傳輸通道是無法滿足這樣的需求，即便是改成 64QAM 的調變後整個 Payload 仍嫌不足，此時若調動 GI 參數，則會造成單頻網條件變更，要求鄰近兩發射站間距離變短，而造成建置成本投資變大，若調動 CR 參數則可能會造成單頻網的同頻干擾區域變大，造成接收用戶接收困難。

所以在經過相關單位測試及會議商議後，經濟部於 2005 年頒發台灣「地面數位電視接收機基本技術規範」中定義，**高畫質電視接收必須以 MPEG4/H.264 格式進行接收解碼**，這也間接的規範了台灣地區 DVB-T 高畫質電視必須改以 MPEG4 /H.264 壓縮格式來進行傳輸使用。而 MPEG4 /H.264 壓縮一個高畫質電視內容所需的資料量則視內容的複雜程度而有所不同，英國衛星傳輸公司 SkyHD 提出通則性的需求建議如表格 1 所列。

表格 1、英國 SkyHD 在 H.264 廣播用影像壓縮率提出的建議值

HDTV 內容	新聞/卡通	一般性節目	運動節目
Bitrate (bps)	8~10 M bps	10~12M bps	15~20M bps

(資料來源：IBE 月刊，September/October 2006)

自 2005 年開始公共電視接受新聞局委託，進行高畫質電視頻道的建置規劃，並由 NCC 於 2008 年核發 Ch.30 做為高畫質測試頻道使用，為期一年。這項測試計畫主要目的除了對 HDTV 傳輸標準的評估及相關服務運作等測試外，對 HDTV H.264 壓縮/解壓縮的影音解析度評估，必須做更深入的研究分析。公視在目前 HDTV 測試計畫中以 DVB 組織所定義之規範進行及發展，相關規範依循 DVB 組織所公佈之 HDTV 標準及介面協定，且依據台灣地區高畫質電視採 AVC H.264 格式進行壓縮編/解碼等規範進行高畫質電視廣播。

2.2. 以固定接收做為高畫質電視網路條件下的規劃

由於發射網路調變參數的不同，也會有資料量及接收電場強度的不同因素相互牽動，我們在表格 2 及表格 3 中可以清楚的了解到其數位電視網路建置的特性與限制。

表格 2、不同調變參數所提供的資料傳輸量

Bandwidth	Modulation	GI	CR	Payload
6Mhz	16QAM	1/4	2/3	9.9M bps
6Mhz	64QAM	1/4	2/3	14.9Mbps

表格 3、室外固定接收最低電場強度需求 (單位：dBuV/m)

Modulation IFFT=8K Code Rate 2/3 Guard Interval 1/4	通道模型	16-QAM	64-QAM
	Gaussian	35.9	41.7
	Ricean	37.7	45.1
	Rayleigh	43.3	50.5

由於高畫質測試頻道規劃主軸是以固定接收為主，所以其網路通道便有兩種 16QAM 或是 64QAM 調變模式可做為選擇，這也是整個頻道測試及營運評估的重點。這兩種調變模式提供不同的資料承載量，決定了高畫質節目受資料壓縮率限制的條件。

DVB-T 網路在固定接收的條件下，我們由表格 2 及表格 3 所示。其 Payload 分別為 9.9M 與 14.9M bps；而接收場強經計算後最低門檻分別為 43.3 dBuV/m 和 50.5 dBuV/m。在這 2 種組合中，如果考慮高畫質電視是以一般家庭使用者為考量，也就是固定接收為主，則高畫質電視節目頻道以 MPEG4/H.264 壓縮後所需的最低傳輸資料量將會是決定網路建置的主要決定因素之一。

2.3. DVB-T 通道實現理想中的中高資料流量的固定接收形式

2.3.1. 是否以 64-QAM 為主作為高畫質電視固定接收的調變模式？

以目前台灣 DVB-T 工程較能實現組合的 2 種是中高資料流量作為固定接收使用，在保留一些必要的頻寬外，其應用在影像壓縮的有效位元速率約為 8~13 M bps，如表格 4 說明；而在這 2 種模式中，16QAM 雖可以提高場強覆蓋率，但卻會降低資料傳輸率為代價，然而我們經過電腦場強的模擬驗證可以得知，以 64QAM 的模式在固定接收的條件下，全省 10 大主要轉播站的覆蓋面積覆蓋人口，與 16QAM 的差異僅為 0.8%，如表格 5，所以初期是以 64QAM 作為規劃高畫質電視的網路調變模式為主。

表格 4、種調變模式下影像有效位元速率

調變模式	Total Payload	PSI/SI table	Data (for EPG/ SSU/ Sub-title)	bitrate	
				Video	Audio
16QAM	9.9M bps	500k bps	≐1M bps	8M bps	400k bps
64QAM	14.9 bps	500k bps	≐1M bps	13Mbps	400k bps

表格 5、全省 10 大站場強模擬計算面積及人口值

10 大轉播站	總涵蓋面積(KM2)	總涵蓋人口	占總人口%
16QAM (Threshold = 43 dBuV/M)	26499.69	21532848	95.21%
64QAM (Threshold = 50 dBuV/M)	22362.88	21351142	94.41%

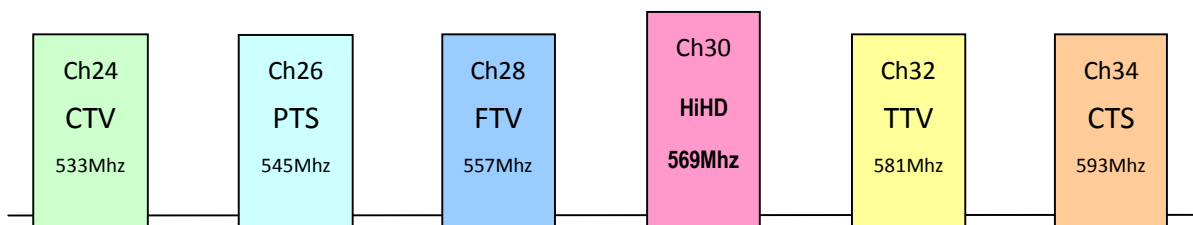
1. 高畫質測試頻道計畫中全台灣 10 大轉播站(竹子山、枕頭山、中寮、花蓮、台東、五指山、三義、南投、宜蘭、壽山)。
2. 電腦以 Fesnel model 模擬分析 10 大站電波場強覆蓋計算。
3. 人口統計來源：2004 年內政部統計人口總數。

2.3.2. 將來是否會改用 16QAM 調變模式來提高室內覆蓋率?

英國 1998 年開始進行地面無線數位電視廣播時，採用 64QAM 調變(目前少數服務中仍採用)。但 OnDigital 的“付費收看”服務開始一年半後卻宣佈失敗，主要原因歸咎於電波覆蓋率不足。因此，當 BBC 和 Castle 獲得 OFCOM 授權合作而重新啟動免費收看服務 Freeview 時，大部分地區調變模式改為 16QAM 調製，它可使現有電波覆蓋範圍的覆蓋率由 66% 上升到 83%。

這是英國在整個網路建置的過渡時期，數位發射功率受到對現有地面類比電視廣播鄰近頻道或同頻道的干擾之限制而採取的補救措施，然而降低接收門檻值而增加在原有覆蓋範圍內的覆蓋率所付出的代價是頻道內資料傳輸速率由 24 M bps 下降為 18 M bps(由 5 套 SDTV 服務減少為 4 套)。

然而這種情況基本上應該不至於發生在台灣，因為我們目前現有的數位電視頻道(包含高畫質測試頻道)都規劃集中在 Ch.24~34 之間，並未有數位與類比頻道鄰近的情況發生，如圖表 1。除非是要以 16QAM 的調變模式來降低接收的困難度，兼顧到室內接收(含可攜式接收)和行動接收，而 16QAM 的模式也會減少其資料通道的承載量，所以這一個前提也必須是 HDTV 頻道影像的資料壓縮量能降到 8M bps (且經評估畫面品質在一定品質以上)才有可能實施。



圖表 1、我國數位電視頻譜分配示意圖

2.4. 小結

- 1) 採用 DVB-T 的 64QAM 調變模式作為接收端的固定接收時，可在原定覆蓋區域內提供無線網路傳輸及提供較高的資料承載量，可用來承載至少一個高畫質頻道內容服務。
- 2) DVB-T 的 16QAM 調變模式作為接收端的接收條件時，雖有利於同時兼顧室內/外、可攜式接收、行動接收等接收條件，但相對的頻寬所能提供的資料承載量相較於 64QAM 則將損失至少約 1/3。
- 3) 目前許多頭端設備製造商不斷改進影音即時壓縮的效率，以 H.264 格式進行高畫質電視內容壓縮技術更是精進，而在通道承載量及內容壓縮率兩者間如何取得一個平衡點將會是日後持續的挑戰。

參、影像主/客觀評量測試方法

3.1. 說明

數位電視系統為了傳輸通道頻寬及儲存媒體容量等等限制而必須進行資料壓縮，充分的利用了人眼的視覺特性，包含眾多的破壞性壓縮演算方法。影像壓縮可以看作一種傳輸資料量、影像品質視覺感知失真和演算法複雜程度之間的一種折衷與妥協。在傳輸過程中信號也會有所損失。因此影像直覺觀賞品質的好壞是衡量系統的重要指標。雖然數位電視系統與類比電視系統的標準不同，但兩者的最後目的卻是相同的，以視覺特性為根據提供影像訊息，系統測量應考量與視覺特性等相關因素，這與國際公認的影像品質評量以主觀評量方法是一致的。

目前影像品質評量方法可分為主觀評量和客觀評量兩種。而主觀評量方法大都採用國際電信聯盟 ITU-R BT. 500-11 (2002 年)和 ITU-R BT. 710-4(1998 年)建議書之精神而定義。由於雙刺激連續品質評量法(The double-stimulus continuous quality-scale method- DSCQS)具有靈敏度高、穩定性好的特點，較適合用於高畫質電視系統影像品質的主觀評量，所以我們將運用這個評量方法來進行 HDTV 相關的測試與評估。

3.2. 影像品質評估方法

由於影像品質的評估包括主觀評量和客觀評量兩個部分，這兩項指標同等重要。考核項目主要是以下幾個項目：

- 影像解析度

數位影像的解析度要考量水平和垂直兩個方面，水平和垂直解析度對評量數位影像品質同等重要。

- 影像圖幀率

數位影像的圖幀率測試不能只是簡單採用幾個圖幀的方法來進行，但目前這次數位電視影像的幀率測試基本上還是以每秒 29.97 frame 為主。

- 影像壓縮比

由於一般的測試系統的影像壓縮比是可以調整的，因此測試的方法計畫是以固定的幾組壓縮參數下進行評量，但應避免在低壓縮比下測試，而在實際應用中卻設定為高壓縮比。

- 影像品質評量

主要評量影像還原品質的真實程度和等級。

3.3. 影像品質的主觀評量

影像主觀評量就是經由人來進行觀察影像，對影像的優劣作主觀的評定，然後對評分進行統計平均，而所得出的評量結果。這評量出的影像品質與評鑑人員的特性及當時的觀察條件等等因素有關。

3.3.1. 主觀評量的分類

通常，主觀評量分成兩類。第一類是確定系統在最佳條件(及理想狀況)下的性能的評量，即所謂的品質評量。第二類是確立因發射傳輸而導致的非最佳條件下維持影像品質能力的評量，即所謂的失真劣化的評量。

3.3.2. 主觀評量的實施

為確保主觀評量在統計上有意義，選擇評量受測人員時既要考慮有未經過相關訓練的一般觀察人員參與，又要有對影像技術有相關經驗的專業觀察人員。另外參與評量的觀察人員至少要有 15 人以上。測試的條件儘可能與使用條件相當。

在實驗心理學的經驗中，當測試感覺時，相對評量比絕對評量的結果更加穩定、準確。在 ITU-R BT.500 建議的許多評量方法中，由於雙刺激方法採用了基準影像比對，其結果比單刺激方法具有更高的靈敏度和穩定性，特別適合於對高品質電視系統或設備性能的主觀評價。因此，在 ITU-R BT.710 (HDTV 影像品質的主觀評價方法)建議將雙刺激連續品質標準方法和雙刺激失真劣化標準方法作為 HDTV 影像品質和發射系統劣化的主觀評價方法，並分別以 HDTV 原拍攝影像和未劣化的播出影帶作比對基準。

- 雙刺激連續品質標準方法(DSCQS:Double Stimulus Continuous Quality Scale)

在這方法中，需要對每個測試影像的兩種狀態進行評分。其中一個是來自原始信號源的影像，即基準影像信號；另一個是經過被測系統壓縮/解壓縮後輸出的影像，即測試影像。基準和受測影像交替顯示兩次之後由受測人員進行評分。不同測試影像的一連串顯示評分過程中，基準和測試影像呈現的先後次序以假的隨機方式變動(受測人員事先並不知道哪一個是基準影像)，要求受測人員對每對影像的整體品質進行評分，並在評分表上作出標記。

評分表由若干對縱向標度線組成，以對應到每個測試影像兩種狀態的評分表。為避免量化誤差，標度線提供連續標度，且被分成 5 個等級，如表格 6(a)相當於標準的 5 級品質刻度範圍。在表的左邊標有與不同等級相對應的品質術語，即「優、佳、尚可、差、劣」作為受測人員評分的依據。

- 雙刺激失真劣化標準方法 (DSIS:Double Stimulus Impairment Scale)

這方法適用於非最佳條件下對系統保持品質性能的評量，它是評鑑人員依據基準影像(無失真劣化影像)比對被測影像(有發生失真劣化)作出評量的方法，並使用五級的失真劣化標準表。

在雙刺激失真劣化標準方法(DSIS)中，首先顯示來自信號源無損傷的影像，即基準影像；然後顯示可能是經過被測系統後輸出的影像，即測試影像。基準和測試影像交替顯示一次或兩次之後，要求受測人員根據無損傷的基準影像，對被測影像品質的受損情況進行評分，並在評分表上作出標記。在受測影像或片段中也包括無損傷的影像或片段。

雙刺激失真劣化標準方法使用 5 級劣化評分刻度。其評分等級和對應的劣化術語如表格 6(b)所示，要求受測人員根據被測影像品質的劣化情況，選用適當的評分等級來進行評分。

由於人類的心理和生理等狀況會因人而異，受測人員對同一影像的品質評量也會有所差異，所以只有經由統計理論進行大量的人次進行評量後進行平均分析，如此才能得到按照一定的概率分佈的數據。

3.3.3. 影像品質主觀評量方法

在影像品質的主觀評量方法中又分為兩種評量計分方法，就是國際上通用的5級評分的品質等級和失真劣化等級，如表格6所示。他是由評鑑人員根據自己的經驗，提供一組標準影像作為參考基準，幫助評鑑人員對影像品質作出合宜的評量。一般來說，對非專業人員多採用品質標準等級，對專業人員則使用失真劣化標準等級為宜。

表格 6、兩種評量的影像5級評分表

(a)DSCQS 評分表

品質等級	評分
優	5
佳	4
尚可	3
差	2
劣	1

(b)DSIS 評分表

失真劣化等級	評分
沒察覺	5
有察覺但不討厭	4
稍有討厭	3
很厭惡	2
劣	1

主觀評量方面，應主要從以下幾方面進行評量：

- 馬賽克效應：還原影像中對於單色區域畫面還原存在的色塊呈現。
- 邊緣處理：對於影像中的物體邊界還原的真實程度和線條（橫、豎、斜方向）還原的真實程度，主要觀察邊界的對比度和變形情況。
- 漸層顏色平滑度：還原影像中對於單色區域畫面還原後的顏色層次豐富程度。
- 畫面的還原性：指畫面還原的真實性，包括畫面的完整性、是否存在色差、對還原影像的整體接受程度。

（以上四項在對參考影像的全部內容測試後做出綜合評量結果）

- 快速運動影像處理：觀察快速運動參考來源下影像還原的連續性。
- 低照度高雜訊環境影像處理：觀察低照度環境影像還原的清晰度。

因為若不在嚴苛的條件狀態下是無法測試出真實的評量數據，所以測試素材應該是經過評估篩選後，其內容是"嚴格的，但又不過分"，即包含對各種評量因素，如靜止畫面解析度、動態畫面解析度、亮度重現、影像顏色重現和運動畫面重現等具有比較敏感特點的那些內容。而測試素材同時又應包含能代表電視節目的典型內容。評量應該採用專業人士和非專業人士共同組成的專家組來評量，對每個專案按照5分制打分。

3.4. 影像品質客觀評量說明

影像品質的客觀評量方法主要是考核影像還原的解析度和信雜比。就目前而言，由於台灣沒有相應的標準，所以從技術上現在還沒有一種有效的方法來測試準確

的資料，根據現在的條件切實可行的方法就是製作參考視頻原始來源。

參考視頻原始來源的製作主要考慮數位影像壓縮的特點、現場節目內容表現特點及在製作的時候可以選擇不同節目的特點。

對電視系統而言，最真實的測量應當是評鑑人員的滿意程度。測量靜態測試信號的失真可以間接地表示出類比信號和數位電視視頻的影像品質特性。但是，表現壓縮電視系統的影像品質十分困難。壓縮系統中的影像品質會隨著壓縮資料速率、影像的複雜程度以及所採用的編碼演算法等條件不同而有所改變。有鑑於靜態測試信號的特性，它無法提供壓縮視頻系統中影像品質的真實特性，對於壓縮視頻系統中影像品質的測試，雖然可以使用具有自然內容和動態的測試場景，通過人眼觀察場景以報告影像品質評估結果。然而這種評估壓縮視頻系統的方法效率低，也非常不客觀。對於 Codec 設計與進行品質的評估而言，憑藉人眼觀察的主觀測試方法是不實用的，這就要求建立影像的客觀評量的方法。

影像的客觀評量利用數學模型測量影像品質，與主觀評量相比具有較多的優點，因此實用性較高。

傳統的數位影像品質客觀評量方法是在於數據傳輸過程中均方信雜比的想法，代表性的方法有 MSE(均方誤差值)、PSNR(峰值信雜比)、WMSE(加權均方根差)等等。

- PSNR(峰值信雜比)方法

衡量數位視頻最常用的客觀測試指標是 PSNR。按相關標準，在測試狀態下，使用符合標準規定的多種測試信號和圖卡，對數位電視信號進行 8 bit 量化，各像素雜訊值 $N(x,y)$ 指的是解碼重建影像各像素信號值 $Y(x,y)$ 與來源影像對映像素信號值 $X(x,y)$ 之差。把各像素的雜訊之和對整個圖禎影像像素總和(M)求平均得到 PSNR。PSNR 雖然是一個客觀值，且便於測定，但是視覺是一個非常複雜的系統，既有生理方面的因素，又有心理方面的因素，PSNR 值小，未必視覺就很滿意。例如影像呈現壞點，雖然很醒目，主觀評量很差，但把這種局部失真劣化平均到整個影像，計算出的 PSNR 值有可能仍然很小。

PSNR 是用來在對數級上描述品質，並且依賴於原始信號和改變後信號的均方差 (MSE)：

$$PSNR_{dB} = 10 \log_{10} \frac{(2^n - 1)^2}{MSE}$$

PSNR 可以很方便而快速地被計算出來，這樣它就成爲了一種很常被使用的方法，並用來測量壓縮和解碼後影像/影像的品質。

PSNR 方法有幾個限制，PSNR 需要一個原始的影像做爲對比，但是這也許是無法在所有情況下都可以實現的，也難保所謂的原始影像沒受過影響。PSNR 不能準確地給出主觀的影像品質數值。對於給定的一個影像或一個影像片段來說，高的 PSNR 通常說明品質好，低 PSNR 說明品質差。然而，一個特定的 PSNR 值並不代表等於絕對的主觀的品質。主觀上感覺好的影像不一定 PSNR 值高。這種情況下，是人類眼睛的觀察敏感區所接收到讓人感覺清晰度很好的

結果，但信雜比卻不一定高。

3.5. 主觀評量實施依循規範

3.5.1. 引用標準

3.5.1.1. ITU-R BT. 500-11 Methodology for the subjective assessment of the quality of television pictures (2002年)

3.5.1.2. ITU-R BT.710-4 Subjective Assessment Methods for Image Quality in High-Definition Television (1998 年)

3.5.2. 名詞解釋：(主觀評量測試採用下列定義)

3.5.2.1. 主觀評量 subjective assessment

直接利用受測人員對被測系統品質的主觀反應，來確定電視系統性能的一種方法。

3.5.2.2. 品質評量 quality assessment

在最佳條件〈即理想收看條件〉下，確定電視系統性能的主觀評量。

3.5.2.3. 失真劣化評量 impairment assessment

在非最佳條件〈如發射和傳輸/接收等條件〉下，確定系統保持品質性能的主觀評量。

3.5.2.4. 測試影像 test picture

用於電視影像品質主觀評量的、在影像素材上有特定要求的影像或影像序列。

3.5.2.5. 評分標準 scale

在主觀評量中使用的評分尺度。

3.5.2.6. 標準解析度電視 standard definition television (SDTV)

其等效品質基準是由取於 ITU-R 601 建議書中 4:2:2 取樣的影像，經過資料碼率壓縮編/解碼處理後，而其主觀評量品質優於一般的標準解析度電視的數位電視。

3.5.2.7. 高解析度電視 high definition television(HDTV)

高解析度電視在水平和垂直兩個方向上，畫面解析約為標準解析度電視的兩倍，並具有 16:9 的長寬比例。

3.5.3. 主觀評量的一般要求

3.5.3.1. 受測條件

數位電視影像品質的主觀評量依據 ITU-R BT.710 中建議，如

表格 7 中定義的觀看條件下進行。

表格 7、數位電視影像品質主觀評量觀看條件要求

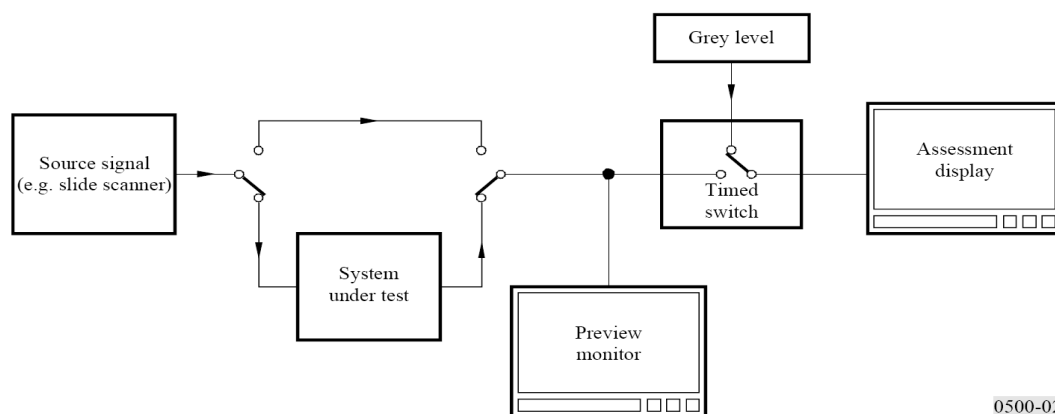
項目	HDTV 參數值
(1) 觀看距離	3 倍的影像高度
(2) 顯示螢幕的最大亮度	150~250 cd/m ²
(3) CRT 電子束流截止時，螢幕亮度與峰值亮度之比	≤0.02
(4) 暗室中，黑階亮度與最大亮度之比	約 0.01
(5) 顯示器背景亮度與最大亮度之比	約 0.15
(6) 室內環境燈光照明	宜低
(7) 背景光和照明燈光光源的色溫	D65
(8) 背景光部份對受測人員的視覺角度	高 53 度， 寬 83 度
(9) 受測人員的座位位置角度	水平方向在中垂線±30 度內 (垂直方向待定)
(10) 建議顯示影像(對角線)尺寸	1.4m

備註：

1. 條件(2)螢幕的最大亮度對應於 100%的視角信號峰值。在 HDTV 的最大亮度達不到規定值之前，應滿足 $\geq 70 \text{ cd/m}^2$
2. 條件(3)受室內照明和顯示對比度範圍的影響。
3. 條件(4)黑階電位對應於 0%的視頻信號峰值。
4. 為了滿足條件(3)和(5)，應該設置室內照明。
5. 若 HDTV 達不到條件(8)規定值之前，應滿足高 ≥ 28 度，寬 ≥ 48 度。
6. 若 HDTV 達不到條件(10)規定的顯示尺寸之前，應滿足 $\geq 76.2\text{cm}$ 。

3.6. 主觀評量系統說明

主觀評量系統設備應該是高畫質電視 HDTV 設備，不影響被測系統或設備的畫面品質，系統的理想方塊圖如圖表 2 所示。



0500-02

圖表 2、系統的原理方塊圖

受測人員從評量用監視器上看到的是經過定時開關送出的 HDTV 影像。定時開關的輸出信號或直接來自信號源，或已經過被測系統，或來自灰階信號產生器。灰階信號產生器的輸出設定為 28.6% 的白色峰值視頻電位。

3.6.1. 測試信號源

測試信號源是直接提供受測試的基準影像，輸入後用來進行受測系統的測試，它應該使用的是電視標準中最佳品質。對於整個評量的環境中若缺乏對映的基準信號的弱點比較，將決定是否可以獲得確定結果的重要因素之一。

數位式儲存的影像和影音片段播放器是最可靠的再生的信號來源，因此這種方式將是最常用的型態。他們可以在實驗室之間交換使用，使評量系統比較更有意義。錄影帶或是電腦硬碟機是最通用的形式。

35 毫米影片掃描器亦可以提供一個可靠的靜態影像的來源。可得到的結果足夠對傳統標準解析度電視的評估。電影影片的颜色和其他特性可以把一個不同的主觀的表現作為攝影棚內攝影機的畫面。如果這會影響結果，那就直接將攝影棚內攝影機的影像來源拿來使用，雖然這可能沒那麼方便。通常，影片掃描器可以透過調整來符合最主觀的影像品質檢測所需的。

而顏色圖板(Color-matte)最常用來檢測串流的處理能力。在攝影棚作業時，顏色圖板(Color-matte)對於攝影棚內燈光的表現是相當的敏感。因此評量應該儘可能地使用一些特別的顏色圖板(Color-matte)，這將會得到高畫質的結果表現。如果有需要，可將移動的物件放在前景滑動。

這將經常被要求考慮這種測試的模式，在早期的信號處理中，不論在哪一種系統測試下的性能都可能受某些處理的影響的而造成結果改變。因此無論何時的測試，在可能處理劣化衰減的環節中的部分上進行測試，雖然不見得是合乎需要或是合理，所以導致的信號處理程序應該被透明記錄，適合以後試驗使可再提供為

資料串流，當想檢查劣化衰減可能是由一連串處理的環節及流程而累積造成。這樣的紀錄該被放在試驗材料的圖書館中而可被查詢及調閱，滿足未來使用者之需要，這包括他們一個被記錄的信號的歷史的詳細的陳述說明。

3.6.2. 測試素材的選擇

一套測試影像通常包含一定數目的靜止影像和含運動物體的影像片段，每個片段大約持續 10 -15 秒鐘。

測試素材應該是經過評估篩選後，其素材內容是"嚴格的，但又不過分"，即包含對各種評量因素，如靜止空間解析度、動態空間解析度、亮度重現、影色重現和運動重現等具有比較敏感特點的那些內容。而且，測試素材同時又應包含能代表電視節目的典型內容。

在電視評量要求有許多方法被接受而建立測試素材。而實際上，各種特別的試驗素材應該用來處理個別特殊的評量問題。對典型的評量問題和用來處理這些問題的試驗素材的調查如表格 8。

表格 8、測試素材的選擇

評估問題	使用素材
用一般典型的素材來作系統整體性能評估	一般，具有挑戰性但不是太過分
表現能力，關鍵的應用(處理後的能力表現)	範圍包括應用的非常具有挑戰性的素材進行測試。
系統的適應性能	使用非常嚴苛的素材來進行計畫上的測試。
系統缺點鑑定和可能的改進方式	關鍵及屬性特殊的素材
被測試系統可看見變化的因素	豐富、範圍廣闊的素材
在不同的標準中的變換能力	不同的特性差異(例如：field rate變換)

*全部測試素材應都可能是電視節目內容的一部分。

一些參數可以引起大多數畫面或者片段的畫質劣化衰減。在這樣的情況裡，有非常少部份的畫面或片段可獲得的一些測試結果，所以仍然提供有意義的評估價值。

不過，受測系統經常會受到某一個畫面或者影像順序或內容而發生影響。在這樣的情況裡，將有為整個計畫而規劃測試作業的節目，必須為可能造成劣化衰減的照片或者影像片段內容進行統計分類。沒有辦法預知這傳輸的形式，通常的情況，選擇試驗素材和結果的解釋一定要非常仔細進行。

通常，包括關鍵性的素材是必要的，因為當要解釋結果時，考慮這方面的可能，但是來自非關鍵的素材推斷是不太可能發生的。如果發生地點或者內容片段的順序會影響結果，給系統在測試下所選擇的素材料應該要選擇關鍵性但不是過分地

要求。一部份不過份的素材內容仍然能可成為正常的計畫小時的一部分。而至少有 4 個項目應該如此，在這樣的情況裡能被使用：例如，哪一半是可以明確評判性的，而哪一半是適當而關鍵的因素。

3.6.3. 受測人員

受測人員即應邀參加主觀評量的評分人員。受測人員通常有兩種類型，即專業受測人員和非專業受測人員。一般都由非專業受測人員來進行主觀評量，當需要精確判斷時，可由受過專業訓練的專業受測人員來進行評量和分析。

主觀評量所需受測人員的數目一般不少於 15 人。

整個評量作業至少要有 15 個以上的觀察人員參與。他們應該是一般的人員而非專家，感覺在他們沒直接涉及電視影像品質評量作為他們的日常的工作的一部分，並且不是有經驗的評估人員²。在一個評量作業之前，那些觀察人者應該被檢查符合(或矯正後) Snellen 或 Landolt 圖表所要求的正常視力靈敏度，並使用特別選擇圖表 (例如 Ishihara) 測試通過其正常的辨色測驗。評量人員的人數則依據上述的測試過程中其敏感性、可靠性和受影響的程度大小而有所增減。

一項研究結果顯示在不同的測試環境的實驗發現有系統上的差別，這會發生在由不同的實驗室之間評量而獲得的結果。為了改進一個實驗的敏感性和可靠性，如果它可以由從幾間不同的實驗室來總合統計其結果，這樣的差別將尤其重要。

可能對在不同的實驗室之間的差別的解釋是因由許多不同的非專家的評估人員中而有不同的技術水準。進一步研究需要承擔評估初這個假說的有效性和如何去證明，確定這個因素而提出的變化的數量。不過，在這期間，實驗人員應該對於他們的評估小組的特性包括儘可能更多的細節使這個因素能更進一步的調查進而變得容易。所提出的建議的數據包括：職業種類(例如 廣播業從業人員、大專院校學生、一般辦公室人員等等)，性別和年齡範圍。

3.6.4. 評量的說明與指示

評量人員應該被詳細說明其評量的方法，哪一些類型可能造成品質劣化衰減的因素，分級的方式，順序並作調整。訓練順序證明範圍和失真確定的類型應該被用於說明，假設的影像並不是在試驗內所使用的那些，但是可比較的敏感性。就品質而論其評量，所以品質就可以確定當時由具體視覺感官的屬性組成。

在每個評量測試階段開始時，應該向受測人員詳細、正確地介紹評量方法和存在的品質受損類型，並進行評量示範顯示，示範顯示應該使用正式測試的影像或影音片段以外的影像或影音片段，但應與正式測試中使用的影像或影音片段具有可比對的特性。

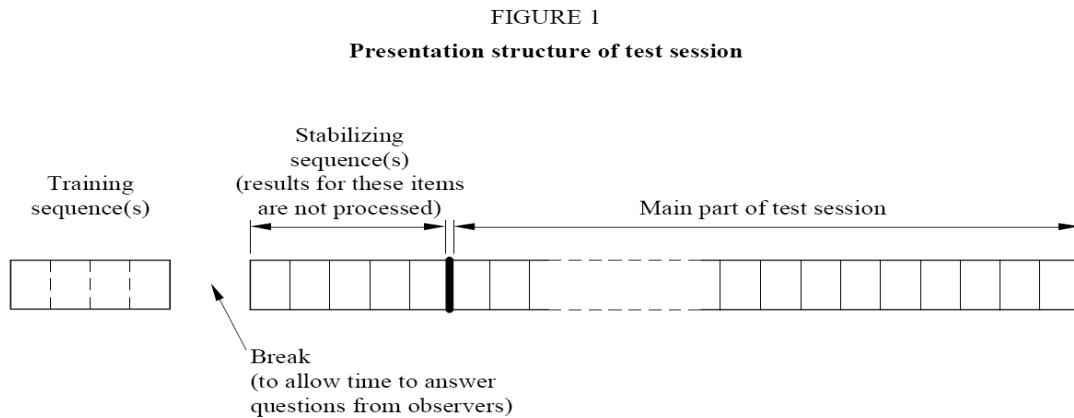
一個評量測試階段包括示範說明在內的時間一般不超過半個小時。在正式測試開始前，需要引入 3 至 5 個測試週期來引導及確定受測人員的判斷力，但該項前導測試結果的資料不納入測試結果的統計中，但必須要考量到的是這些簡介不能引導受測人員的看法而造成某些測試數據的偏差。不同測試影像的顯示順序果用假隨機方式，為了檢測相關性，有些測試可以重複進行，但要避免相同測試影像在

² 初步的結論說明非專家的觀察者可能對高品質影像的傳輸及顯示技術上可以產生更多關鍵的結論。

相繼的測試週期出現。

簡報說明的過程應任意的指派(例如，說明的影像片段的使用)；但是整個測試個過程中應考慮到受測人員的疲累及適應程度，必須在這之間取得平衡點以免造成某些影響。可以在測試過程中一階段到另一階段間重複一些簡報來檢查其測試的連貫性。

測試階段的流程如圖表 3 所示



圖表 3、測試階段的流程

3.6.5. 結果統計分析方式

因為結果會隨測試的範圍而變化，由於大多數評量方法用絕對性的專業術語來解釋判斷是不適當的(例如：一個影像或是影像片段的品質)。

對於每個測試影像或影音片段，應該給出評分值統計分佈的平均值、標準偏差和 95% 可信度區間。

結果還必須包含以下的資訊：

- a) 測試情況的說明。
- b) 測試素材的情況
- c) 影像來源的類型和監視器型號³
- d) 受測人員的人數和類型⁴
- e) 使用的基準系統
- f) 測試總和的平均分數

³ 因為有些經驗顯示出實驗中所使用的顯示器螢幕大小會影響到整個測試結果，所以實驗人員需要在每一次的測試中明確的說明出顯示器的大小，廠牌和型號。

⁴ 有證據在整個工作小組(甚至在非專家小組中)的人員素質水準方面的變化也會影響主觀評量的結果。為了使這個因素變得單純，大部分認為小組成員組成的可能特性及實驗人員的背景資料也必須列入報告的一部份。相關的背景資料可以包括：工作小組的年齡和性別或教育程度及工作的內容等等。

g) 如果淘汰了某些評量人員，要說明原始的和調整後的平均分數和 95%的可信度區間。

3.7. 評分方法和評分標準

3.7.1. 雙刺激連續品質標示法

3.7.1.1. 評分方法

在雙刺激連續品質標度法中，要求受測人員只簡單地對每個測試影像的兩種狀態，即基準狀態〈直接來自信號源〉和被測狀態〈經被測系統〉的影像總體品質進行評量，並在評分表上作出標記。

3.7.1.2. 評分標準

評分表由若干對直向標度線組成，直向標度線成對的排列，以適應對每一個測試影像兩種狀態的評量。直向標度線提供連續標度，且被分成 5 個等分，相當於標準的 ITU-R 中 5 級的品質標度範圍，並在評分表的左邊標有與不同等級相對應的品質術語，即「優」、「佳」、「尚可」、「差」、「劣」、可作為受測人員評分的一般指導，如圖表 4 所示。受測人員記分筆的顏色應與標度印刷線的顏色有區別。

	27	28	29	30	31
Excellent 優	A B	A B	A B	A B	A B
Good 佳	A B	A B	A B	A B	A B
Fair 尚可	A B	A B	A B	A B	A B
Poor 差	A B	A B	A B	A B	A B
Bad 劣	A B	A B	A B	A B	A B

圖表 4、評分表

3.7.1.3. 測試程序

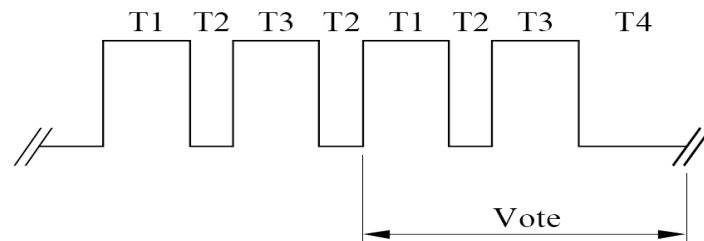
3.7.1.3.1. 示範說明階段

在每個測試階段開始時，要向受測人員詳細、正確地介紹評量方法和評分標準以及影像顯示的順序和時間，但不應出現影響評分結果的暗示。然後進行評分過程的示範，靜止影像和影像片段各示範一次。示範說明應該事先錄音，每次測試開始時播放，以免每次口頭說明有出入或有暗示。

3.7.1.3.2. 影像畫面顯示和評分階段

一個測試週期是對一個測試影像〈或影音片段〉進行評分操作的過程，是組成評量測試階段的基本單元。它包括顯示影像和評分兩部分，對於靜止影像，其基準

影像和被測影像交替顯示 5 次，每次各持續 3-4 秒鐘，最後兩次為評分期間；對於運動影像，其基準影像與被測影像交替顯示 2 次，每次各持續約 10 秒鐘，然後評分。影音片段顯示評分的順序如圖表 5 所示。在不同測試影像的一連串顯示評分過程中，基準影像與被測影像的先後次序以假隨機方式變動。要求受測人員將 T1 顯示影像的評分和 T3 顯示影像的評分，分別記錄在評分表相應編號的某對標度線 A 和 B 上。



Phases of presentation:

T1 =	10 s	Test sequence A
T2 =	3 s	Mid-grey produced by a video level of around 200 mV
T3 =	10 s	Test sequence B
T4 =	5-11 s	Mid-grey

圖表 5、影音片段顯示評分的順序

3.8. 主觀評量的數據統計方法

3.8.1. 雙刺激連續品質標準測試方法的數據統計

首先要確認評量過程的正確性和資料的有效性。先將評分表上所有基準和被測狀態的測量值轉換成標準可分為 1 到 5 的整數，並區分整理出每個測試影像基準影像和被測影像的評分數值，然後進行評分的一致性檢驗，即用每位原受測人員對同一影像在同一測試階段中的不同測試週期結出的兩個評分值來核對，如果評分值相差 2 分(含 2 分)以上，則被視作無效而捨去，若某位受測員在同一測試階段給出的有效評分次數低於其總評分次數的 85% 時，取消其全部評分。

最後計算出每個測試影像兩種狀態的平均分、標準偏差和 95% 取信度區間。

$$\text{平均分數 } U = \frac{1}{N} \sum_{i=1}^N u_i$$

$$\text{標準誤差 } S = \sqrt{\sum_{i=1}^N \frac{(U - u_i)^2}{(N-1)}}$$

95%可信度

區間為: $[U - \delta, U + \delta]$

$$\text{其中: } \delta = 1.96 \frac{S}{\sqrt{N}}$$

u_i ——受測人員 i 評分

N ——受測人員數量

3.8.2. 雙刺激劣化衰減標準測試方法的資料統計

首先要確認評量過程的正確性和資料的有效性，即進行評分的一致性檢驗，用每位受測人員在同一測試階段中對同一影像的不同測試週期給出的兩個評分值來核對，如果評分值相差 2 級或 2 級以上，則被視作無效而捨去，若某位受測人員在同一測試階段給出的有效評分次數低於其總評分次數的 85% 時，取消其全部評分。

最後計算出每個被測影像的平均分數、標準偏差和 95% 可信度區間，計算公式見 (4.1)。

一個測試階段中，全體受測人員給出的累加有效評分次數總和低於應有總評分次數的 85% 時，取消該測試階段的全部評分。

3.8.3. 影像影音內容劣化衰減特性說明

依詢著這些程序，受測的系統可以接受到廣泛的節目素材的實驗測試，有某一部份的品質並無法達到沒有劣化衰減的要求。考量到系統的適當架構，我們必須要同時知道哪些節目的素材將會影響到系統的作品質和這些情況下可能會造成品質下降的比例。實際上，在這系統的考量下，什麼是造成節目內容劣化衰減的錯誤特性。如同一些錯誤特性對於一個不能降低其性能的系統是很重要，而這一部份也變的越來越關鍵。舉例來說，可靠穩定的數位系統可以達到並符合大部份高品質的節目素材的要求，但降低一級的話則會超出這個範圍。

3.8.4. 使用其錯誤特性進行多方考量

在提供一可能的計畫測試素材的範圍上來取得受測系統在處理影像的性能，錯誤特性是考慮系統的適當性的一件重要的工具。錯誤特性可能會引用 3 種模式進行：

- 依據設計的定位來符合其服務需求的最佳化參數(畫面的解析度，傳輸速率，頻寬承載能力)
- 考量單一個系統的適當性(在操作期間預期錯誤或失真的發生率和嚴重程度)
- 評估測試制度規劃相關的適當性(比較錯誤特性並且確定哪個系統將更適合使用)。這應當指出，當一種相似的類型在評選制度可以使用相同的臨界的檢

測時，而不同的類型的系統可能會因臨界檢測而有不同的指標及結果。但是，因為錯誤特性最後可能只顯示品質的不同所呈現的水準來作為系統檢測分析，即使當從臨界測試的不同，所得到具體系統的指標，其特性還是可以直接作為比較。

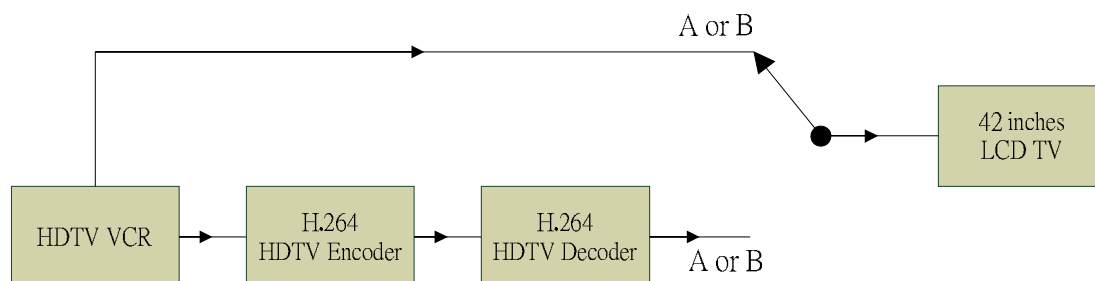
檢測方法在這裡是描述提供一檢測系統特性造成畫面內容錯誤的方法，它可能不完全針對電視服務的一般觀眾所接受。但依循這相關資訊將會更適合將許多觀眾喜愛的節目經由數位系統而進行編碼，並參考他們意見及看法而進行修正的可能性。

數位電視的一個影像內容錯誤特性的例子，在 ITU-R BT1129 中附件 1 敘述其建議。

肆、主觀評量作業程序說明及測試結果分析

4.1. 測試系統規劃：

本次主觀評量採「雙刺激連續品質標準測試方法 DSCQS」進行，其系統規劃如圖表 6。



圖表 6、測試系統方塊圖

4.2. 主觀影像測試方法

4.2.1. 前期作業

- 初選 20 組影像受測影像片段，再經篩選(或經剪輯組合)為 12 組受測影像片段。
- 依據 ITU-R BT.500 與 BT.710 相關規範建議之限制，故僅選擇 10~12 組測試影音片段及 3 種壓縮格式作為本次測試內容。

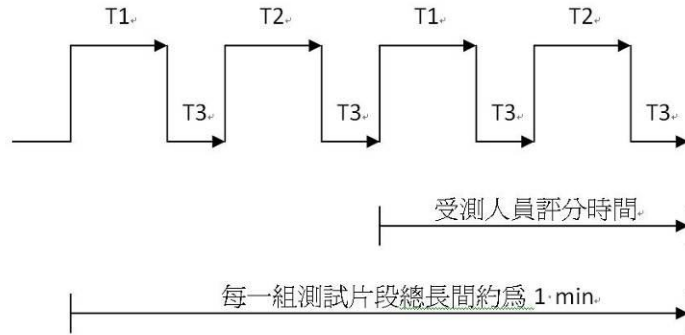
- 主觀影像測試方法目前壓縮率定為：

項次	壓縮格式	壓縮率
1	H.264/ 1920X1080i	6M bps
2	H.264/ 1920X1080i	8M bps
3	H.264/ 1920X1080i	13M bps

- 本次 HDTV 影像壓縮作業採用相關設備如下：

項次	設備名稱	廠牌型號
1	H.264 HD Encoder	Harmonics Electra 7000
2	H.264 HD Decoder	Zinwell ZDT-620HD
3	HD Video recorder	Sony HDW-1800
4	LCD TV	LG 42LG60

- 將原始影像片段經 H.264 壓縮/解壓縮後利用 HDcam 錄影機錄下後再進行剪輯，共計約為 36min 的受測影像片段。
- 每一個受測畫面的循環經剪輯後所構成的測試片段如圖表 7 所示。



T1：原始未壓縮前受測影像畫面，長度為 12sec。

T2：經編解碼後的受測影像畫面，長度為 12sec。

T3：20%的灰階畫面，長度為 3sec。

圖表 7、測試影像畫面的循環順序及時間長度

4.2.2. 測試作業說明

- 測試環境以播放錄影帶於 LG 42 吋 LCD TV 供受測人員進行評分，LG LCD TV 機型為 42LG60，規格說明如附錄。
- 事先必須個別對受測人員進行測試說明及畫面展示，但正式測試時不告知受測人員哪一段畫面是原始片段或壓縮還原畫面。
- 受測人員於評分表上進行評選，評分表依規定分為五個等級，受測人員依序對原始片段或壓縮還原畫面分別(A 組與 B 組畫面)進行勾選，測試完成後須由受測人員填具個人資料後並簽名。
- 依據 ITU-R BT.500 相關規範建議，受測人員接受測試時間一次應不要連續超過 30min，故超過 30min 時應給於受測人員休息 10 分鐘。

4.2.3. 受測人員背景資料說明

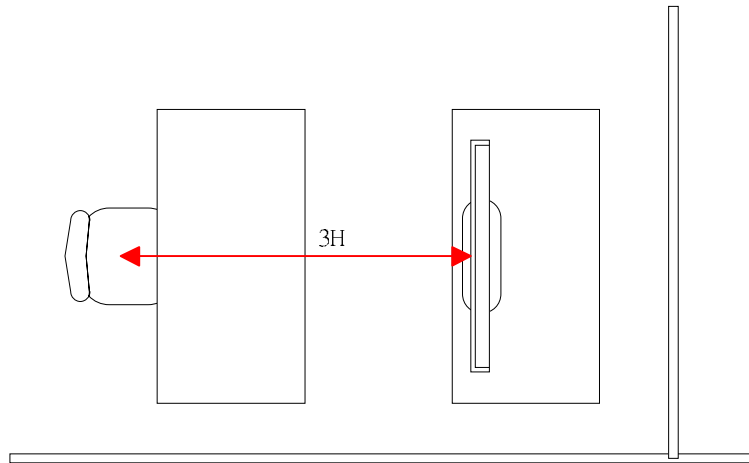
- 受測人員共計有 20 位。
- 人員背景資料如表格 9。

表格 9、受測人員背景資料

性別	男性		女性
	13		7
學經歷	高中	大學/大專	研究所
	1	17	2
年齡	20~29 歲	30~39 歲	40 歲以上
	4	7	9

- 說明：
 1. 依據 ITU-R BT.500 規定，所以受測人員基本上均非受過視訊壓縮相關專業訓練人士組成。
 2. 測試環境依規定擺設外，受測人員於正式評量開始前，並依規定進行 3 段影像片段說明及演練。

3. 測試環境及設備擺設如圖表 8。







※ 觀看距離依 ITU-R BT.710 建議是電視畫面高度(H)的 3 倍



圖表 8、受測環境示意圖

4.2.4. 測試片段內容說明

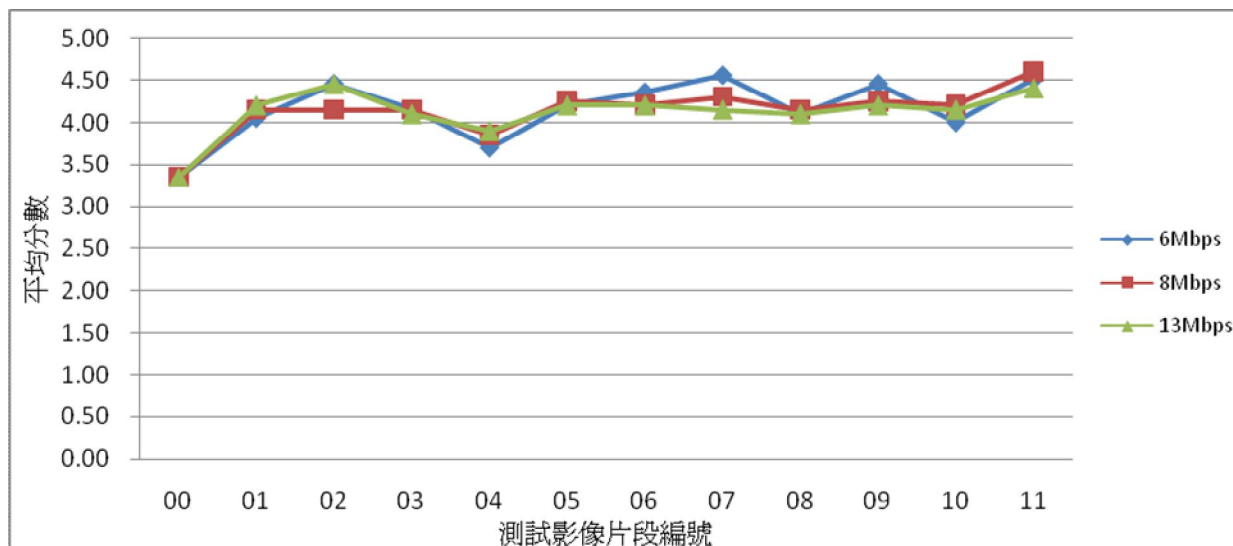
編號	受測影片內容	內容重點說明
00		<p>這段岩石的影像片段是以攝影機由下往上移動拍攝，在原始母帶上就已經因為影像資料大量變化而使得畫面有抖動的現象發生，也利用這現象同時進行壓縮測試，由受測人員進行畫值判定。</p>
01		<p>片尾字幕由下往上流動，背景畫面是海水慢速漂流，造成很多波紋與顏色漸層，這幾個特點測試 H.264 在區塊和顏色的壓縮/解壓縮表現。</p>
02		<p>畫面大逆光及水面浮動光影反射，四人協力划船的線性移動，加上船槳不定時入水造成水紋波動、濺起水花等細節。</p>
03		<p>單人獨木舟在湍急的水流中非預設方向行進，大量的水花及單一色彩的漸層可以考驗壓縮及解壓縮的效率與回復力</p>

04		<p>本片段的處理重點是在於物體旋轉的快速運動及天空藍色的漸層色測試。</p>
05		<p>生態影片中，大量飛鳥運動及攝影機跟拍由右至左移動。可以測試大動態的壓縮現象。</p>
06		<p>賽車片段(1)，車輛高速行進，測試車輛行進運動的畫質變化。</p>
07		<p>賽車片段(2)，車輛高速行進，攝影機移動跟拍，測試不同型態車輛運動的畫質變化。</p>

08		<p>攝影棚內歌曲演唱，重點在於燈光變化，歌唱者臉光不均勻的情況下所造成的畫面損失與壓縮前後差異。</p>
09		<p>運動節目，運用美國職棒節目片段中，測試投球、打擊等動作差異性，另轉場 CG 效果是否會受到壓縮後而有影響，也是測試重點之一。</p>
10		<p>夜間畫面片段，利用暗處高雜訊的特性來作為測試比對。同時也可注意在人群當中，眾多人臉的細節變化及差異。</p>
11		<p>大自然影像片段，植物因風吹動而擺動，壓縮對樹葉間的細節變化；同時也以片段中戶外光影不均勻而使的人臉造成因影後的前後差異。</p>

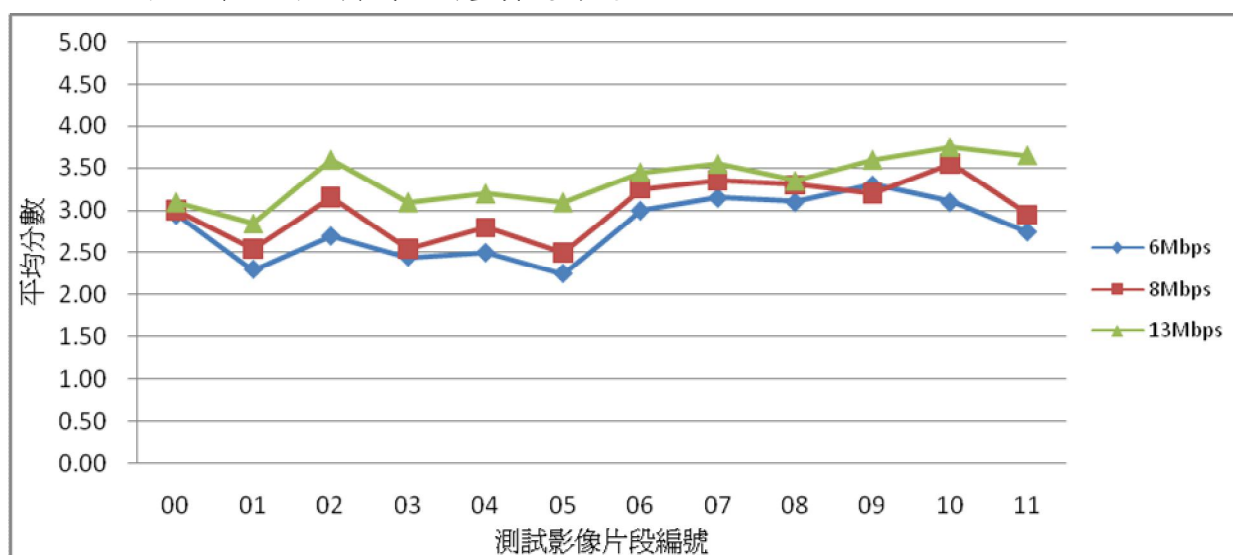
4.3. 測試結果統計分析：

- 由原始畫面品質的評量統計圖(如圖表 9)中可以分析出受測人員對於同一原始未壓縮前的影像片段進行三次評量的結果並無太大的差異，代表受測人員在此次評量的結果是有效的，而且除了00片段外，其於11組片段平均幾乎都在4分(佳)以上。



圖表 9、原始畫面品質的評量統計圖

- 在受測畫面品質評量統計圖(如圖表 10)中各片段的平均分數可以發現，大部分的片段分數都是呈現 13Mbps > 8Mbps > 6Mbps，這也表現出畫面的品質與壓縮率之間的關係，但也並非完全絕對，例如片段 00、08 中即使壓縮率不同，但是平均分數相當接近，可見有些片段即使將壓縮率提高後也無法讓受測人員感受到畫質變得更好。

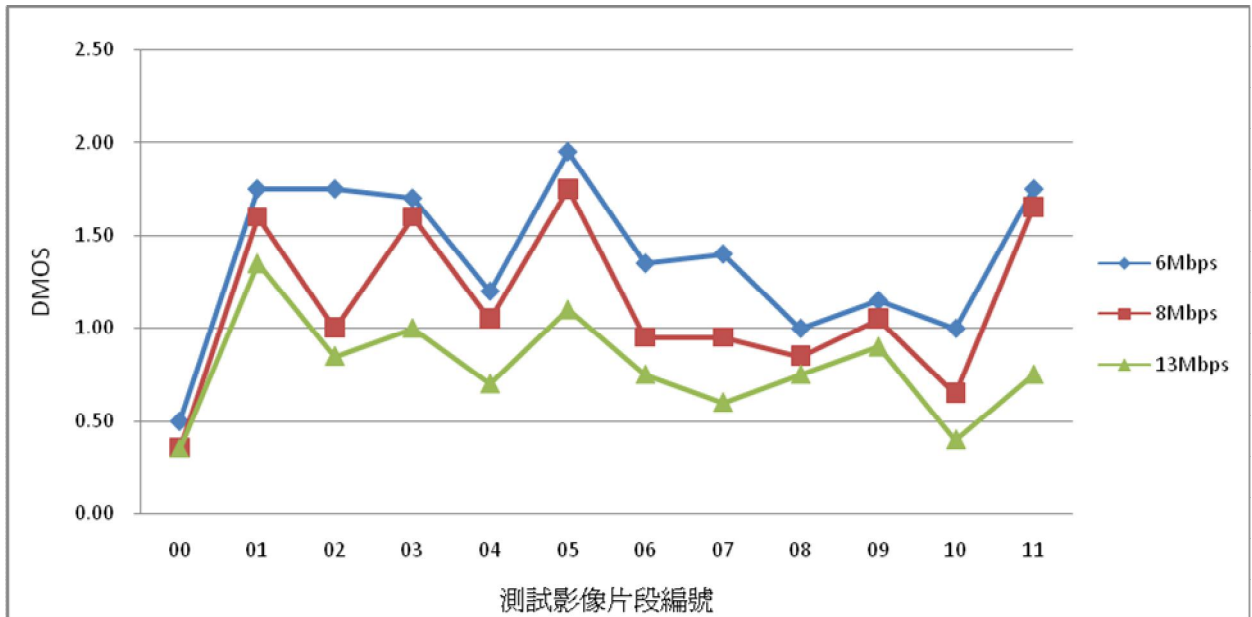


圖表 10、受測畫面品質評量統計圖

- 評量實驗中，讓每位受測人員對觀看原始的影像片段及壓縮還原後片段分別進行評分，取得該項數據後，將這兩者相減，即可得的差值就是表現每

個測試片段經過系統壓縮還原後的損傷劣化程度。這項差值越小則可以表示影像壓縮後的損傷劣化越小，品質與原始影像的差異越少，這數值稱為 DMOS (Difference of Mean Opinion Score)，中文稱為平均主觀分數差值。

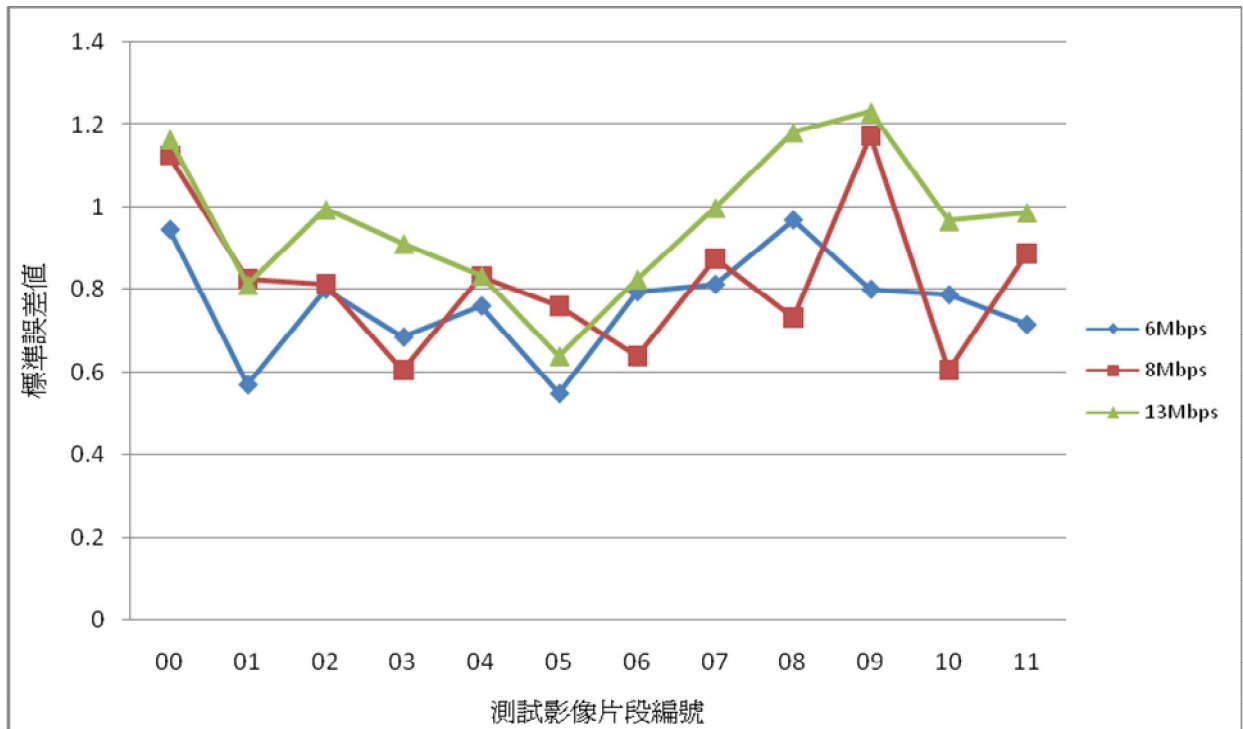
- 在測試結果 DMOS 統計表(如圖表 11)中，整體的數值分析可以發現受測人員對 6M bps 的測試影像片段有較大的差異(平均都有 1~2 的差異值)，而相較於 13M bps 則不容易分辨出原始無壓縮與受測片段之間的差異性(大部分平均都在 1 以下的差異值)。而這也表是以 13M bps 壓縮資料率應該可以滿足大部分在廣播上的影像畫面品質。



圖表 11、評量測試結果 DMOS

- 依(4.1)公式推算出各測試片段的標準誤差值可以發現：
 - 6Mbps 大約是 0.57~0.94 之間，平均為 0.766；
 - 8Mbps 大約是 0.61~1.17 之間，平均為 0.823；
 - 13Mbps 大約是 0.64~1.23 之間，平均為 0.963。

由上述的數據中說明了受測人員對 13M bps 測試片段評分的離散度較大，各個測試片段標準誤差值如圖表 12。這是因為大多數受測人員對於壓縮率較高的影像品質與原始未壓縮影像片段之間品質的差異較小，較難分辨哪一段是原始素材、哪一段個是被評量對象。

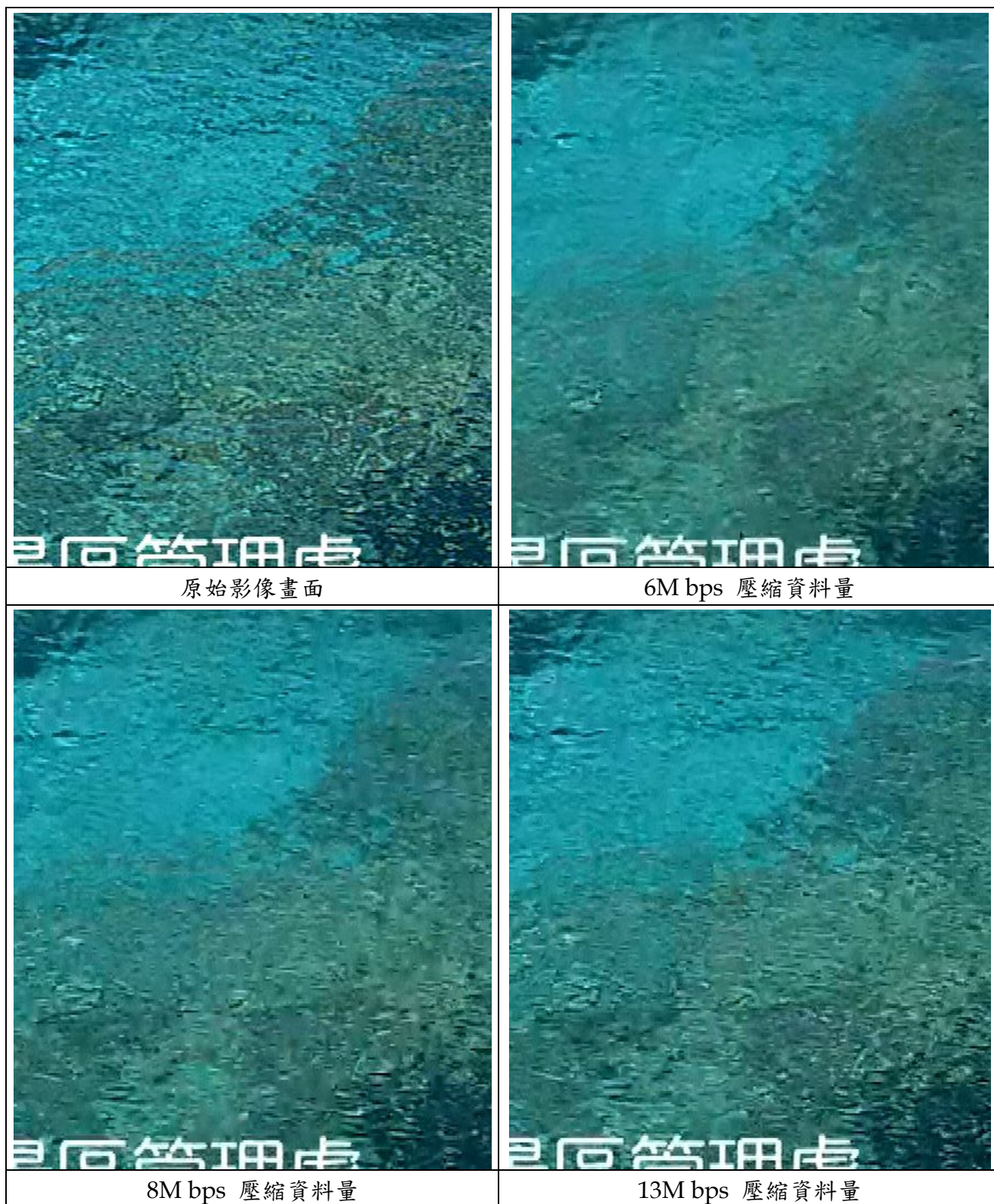


圖表 12、各測試片段的標準誤差值

4.4. 小結

- 在整體的主觀評量測試的統計結果可以看出，H.264 影像視頻編/解碼方案在影像格式為 1920*1080i/59.94Hz、壓縮資料量在 13Mbps 時，受測影像片段的畫面品質與原始素材畫面品質的差別很小，而人眼已不易察覺其兩者之間的差異。
- 可以提出來特別討論的是，我們在測試的 12 組片段中，特意放入一組節目內容因攝影機因由下至上的運動路徑而造成員拍攝帶就已經有抖動的現象(測試片段編號 00)，20 個受測人員評量這段測試時，都判定為畫面品質「尚可」~「差」的等級，而 DMOS 值在 3 種壓縮率下卻都是最低，可見受測人員都認為與原未壓縮畫面幾乎無差別，但畫質卻不好；由此可以得到另一個結論是：如果原拍攝影片畫質已經裂化，所以不論用低或是較高的壓縮率也無法改善其畫面品質，但會容易讓一般人誤解為是因壓縮率不足等傳輸因素而造成畫面品質劣化的原因。
- 我們除了在三種壓縮率進行的主觀評量外，我們在亦同時截取一部分影片中經壓縮/解壓縮後顯示的圖片於報告中供比對(如圖表 13、圖表 14)，可以發現在不同壓縮率之下，畫面之間的差異，包含了影像的解析/模糊程度、區塊大小、邊緣的平滑度等等，而這當然也會直接影響到受測人員在主觀評量上的評等。

圖表 13、實際水波紋及漸層色的差異：由受測畫面中我們發現經 H.264 壓縮之後，所以許多水波紋細節的部份在較低的資料量下，畫面解析度的表現相較於原始畫面會比較模糊，漸層顏色也是較大區塊的顯現。



圖表 13、受測片段三種不同壓縮率畫面比較(1)

圖表 14、人臉邊緣及暗處細節表現：此畫面中因 H.264 以區塊的方式進行壓縮之故，所以在較低的資料量下，臉部雜訊可見較大，畫面解析度的表現相較於原始畫面會比較模糊，臉部的邊緣也較不平整。



圖表 14、受測片段三種不同壓縮率畫面比較(2)

伍、影像品質客觀評量標準與測試結果分析

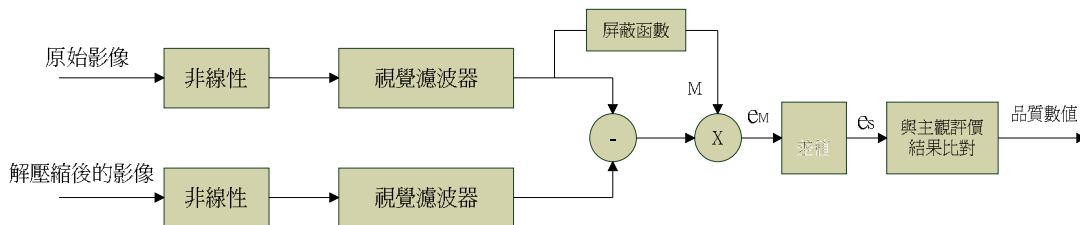
目前對影像品質的評估方法主要分成主觀評量和客觀評量。但很顯然，主觀評量有幾個明顯不足之處：

- 受測的觀察者一般需要是一個比例以上的人數，並且經過訓練用以準確判定主觀評測分，所以需要投入人力和物力投入龐大，測試時程較長。
- 影像內容與情節千變萬化，受測觀察者因有個別差異，容易發生主觀上的偏差。
- 主觀評量無法進行即時監測。
- 因為只有平均分數，如果評測分數低，卻無法確切查證問題在哪裡？

5.1. 客觀評量

客觀評量依據仿人眼視覺模型的原理對影像品質進行客觀評估，並給出客觀評價分數。近幾年來，隨著人們對人眼視覺系統的深入研究，客觀評量的方法和工具不斷被開發出來，其評量結果也與主觀評量較為接近。國際上也成立了ITU-R 影像品質專家群組(ITU-RVQEG: Video Quality Experts Group)專門研究和規範影像品質客觀評量的方法和標準。

VQEG 定義了兩個簡單的技術參數：峰值信雜比(PSNR)和均方差(MSE)。此外，還有許多影像品質模型，這些模型在評量影像品質時都根據人眼視覺特性為主。圖表 15 是一種典型的架構於解碼影像與基準影像差值的影像品質客觀評量模型。



圖表 15、基於解碼影像與基準影像差值的品質模型

該模型的輸入是原始信號和待測的解碼信號，輸出是兩個輸入影像上各圖原幅值之差的和。在整個處理過程中考慮了人眼對影像誤差的主觀感覺特性，以使評量結果與主觀評價所得結果相吻合。模型中的估算考慮了人眼的非線性、視覺濾波器、人眼的屏蔽效應、差值求和。為了使客觀評量與主觀評價結果一致，最後還要使所得的數值範圍和等級描述與主觀測試相對應，對客觀評量的數值進行線性轉換。這個任務是在與主觀評量匹配這一級完成。

影像品質的客觀評量方法分為兩類：“相對評估”(Relative Evaluation)和“絕對評估”(Absolute Evaluation)。

- 相對評估：將處理過的影像（壓縮或經傳輸）與原始影像比較以獲得相對評估的指標值，並根據這些指標值評估影像品質；相對評估一般用原始影片製作時的質量來比對評估，所以準確性高。

- 絕對評估：直接對處理過的影像（壓縮或經傳輸）進行評估以獲得絕對評估的指標值，並根據這些指標值來評估影像質量；絕對評估是以一般線上觀看測試，所以準確性不如相對評估。

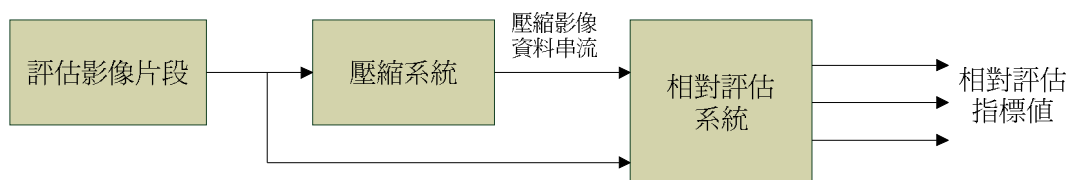
採用客觀評量工具，不僅減少了對人力、物力的需求，而且評量時間大大縮短，甚至可做到即時監測的作業。

以下主要探討在數位電視影像壓縮傳輸應用的信號源頭端和用戶接收端兩種環境下，對影像品質的影響因素以及客觀評量方法。

5.2. 信號源頭端影像品質的客觀評量

信號源頭端的影像品質客觀評量一般採用相對評估，就是在編碼時將壓縮過的影像與原始影像比較，來獲得相應指標的數值，結合多個指標數值對影像品質進行評估。

用於信號源頭端的相對評估模型如圖表 16 所示。



圖表 16、信號源頭的影像品質相對評估

圖表 16 中的相對評估系統可以採用上面提到的架構於解碼影像與基準影像差值的品質模型，也可採用其他模型。

相對評估系統衡量指標按照評估區域分類可分為：空間 (Spatial) 指標和時間 (Temporal) 指標兩種。

- 空間指標：用於考量前後影像圖幀之間比較的指標。
- 時間指標：同時考量連續的兩圖幀或多個影像圖幀時使用的指標。

具體的，主要包括以下指標：

- PSNR：亮度的峰值信雜比，屬空間指標。
- SNR：亮度的信雜比，屬空間指標。
- RMSE：亮度的均方根誤差 (Root Mean Square Error)，屬空間指標。
- 活動能量誤差：額外的能量表明有誤差區塊和雜訊，屬時間指標。
- 重複圖幀：缺少的能量表示有遺漏圖幀，產生影像抖動，屬時間指標。
- 邊緣能量誤差：表是遺漏圖幀或重複圖幀，屬空間域指標。
- 水平和垂直邊緣能量：額外的邊緣能量表明邊緣有雜訊、區塊效應或邊緣雜訊，屬空間域指標。
- 空間頻率誤差：缺少的邊緣能量而出現模糊，屬空間域指標。

5.3. 影像品質的 PSNR 測量

要表現壓縮電視系統的影像品質十分困難，壓縮系統中的影像品質會隨著壓縮資料速率、影像的複雜程度以及所採用的編碼演算法等條件不同而有所改變。有鑑於靜態測試信號的特性，它無法提供壓縮視頻系統中影像品質的真實特性，對於壓縮視頻系統中影像品質的測試，雖然可以使用具有自然內容和動態的測試場景，通過人眼觀察來做為主觀影像品質評估結果。然而這種主觀評估壓縮視頻系統的方法效率低，也十分主觀。對於 Codec 設計與進行品質的評估而言，憑藉人眼觀察的主觀測試方法十分耗時耗力，所以要另建立影像的客觀評量的方法。

影像的客觀評量利用數學模型測量影像品質，與主觀評量相比具有較高的效率，因此實用性較高。

傳統的數位影像品質客觀評量方法是在於數據傳輸過程中均方信雜比的想法，代表性的方法有 MSE(均方誤差值)、PSNR(峰值信雜比)、WMSE(加權均方根差)等等，一般為多數人及設備廠商普遍接受的是 PSNR 的評量方法。

5.3.1. PSNR(峰值信雜比 / Peak Signal-to-Noise Ratio)方法

衡量數位影像最常用的客觀測試指標是 PSNR。按相關標準，在測試狀態下，使用符合標準規定的多種測試信號和圖卡，對數位電視信號進行 8 bit 量化，各像素雜訊值 $N(x,y)$ 指的是解碼重建影像各像素信號值 $Y(x,y)$ 與來源影像對映像素信號值 $X(x,y)$ 之差。把各像素的雜訊之和對整個圖禎影像像素總和 (M) 求平均得到 PSNR。PSNR 雖然是一個客觀值，且便於測定，但是視覺是一個非常複雜的系統，既有生理方面的因素，又有心理方面的因素，PSNR 值小，未必視覺就很滿意。例如影像呈現壞點，雖然很醒目，主觀評量很差，但把這種局部失真劣化平均到整個影像，計算出的 PSNR 值有可能仍然很小。

PSNR 定義為信號功率與雜訊功率之比，通常用分貝表示。在影像處理的過程中，可將信號功率當作是峰值影像取樣值數的平方，而雜訊功率可以為是該影像中誤碼數的平方，即：

$$\text{PSNR} = 10 \log (\hat{S}^2 / \text{RMS}^2) \quad (6.1)$$

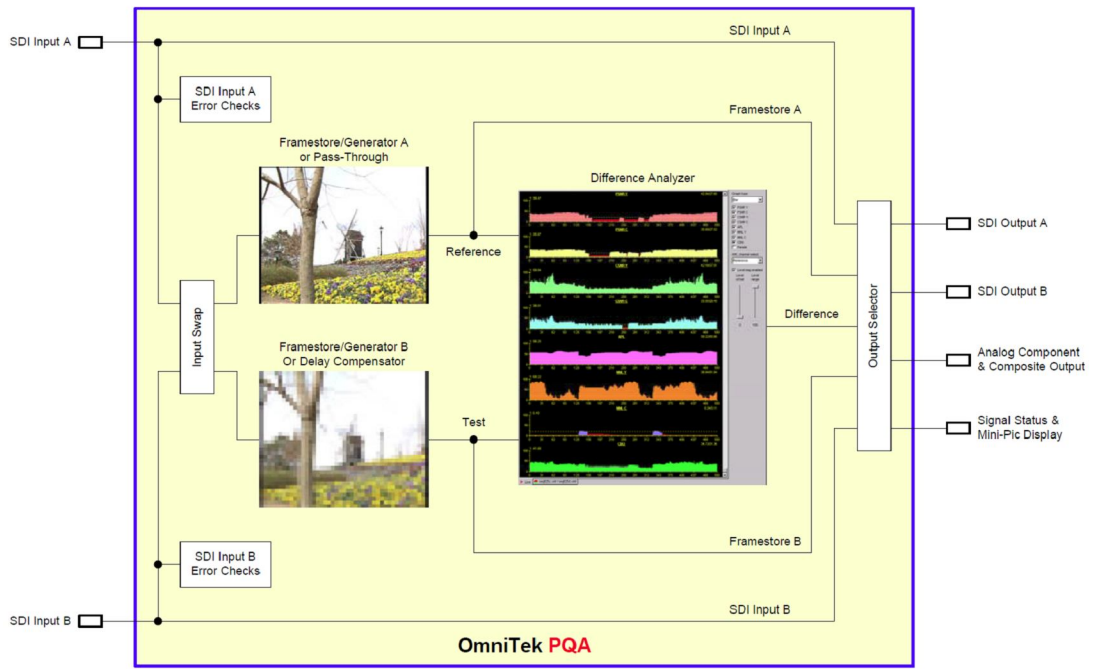
5.4. 測試片段 PSNR 量測值及分析

5.4.1. Omnitek PQA 儀表系統略述

5.4.1.1. 系統架構及功能

Omnitek PQA 在畫質比對的量測上，可以：

- A) 即時進行二組 HD 或 SD 信號輸入進行比對測試。
- B) 即時進行亮度(Luma)與色彩(Chroma)PSNR 值的比對計算。
- C) 畫面邊緣差異比對分析。
- D) 畫面平均亮度的計算。

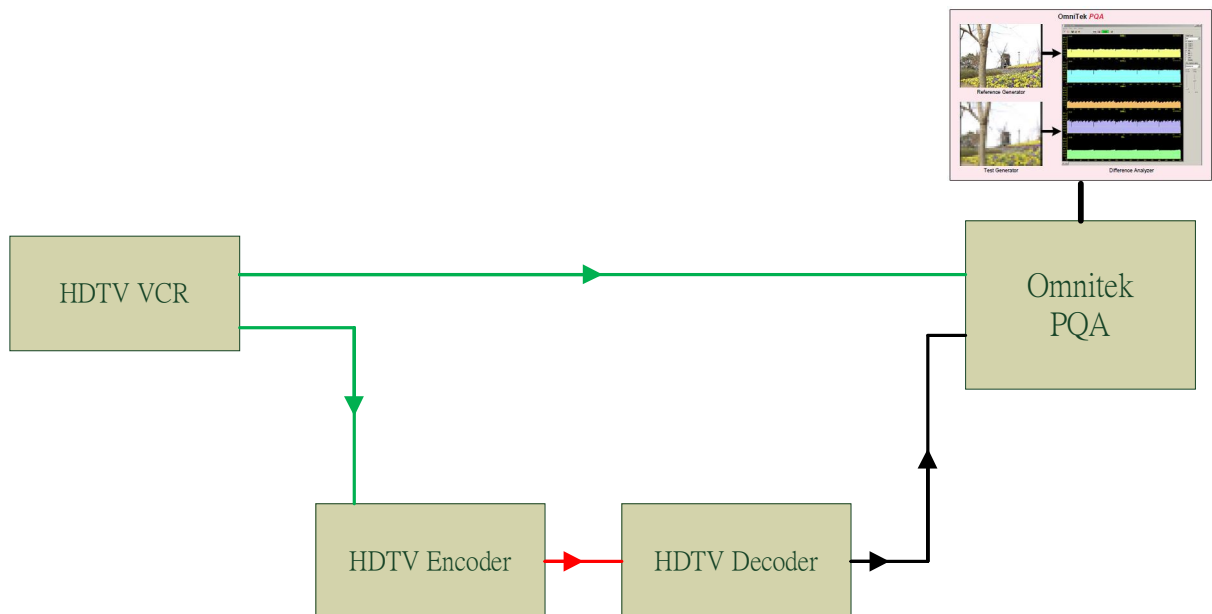


圖表 17、Omnitek PQA 內部信號處理流程示意圖

5.4.1.2. 依循標準規範

依據 ITU-T J.144 的技術規範和來自 VQEG(Video Quality Experts Group)提出的相關技術報告進行畫面 PSNR 的數值計算。

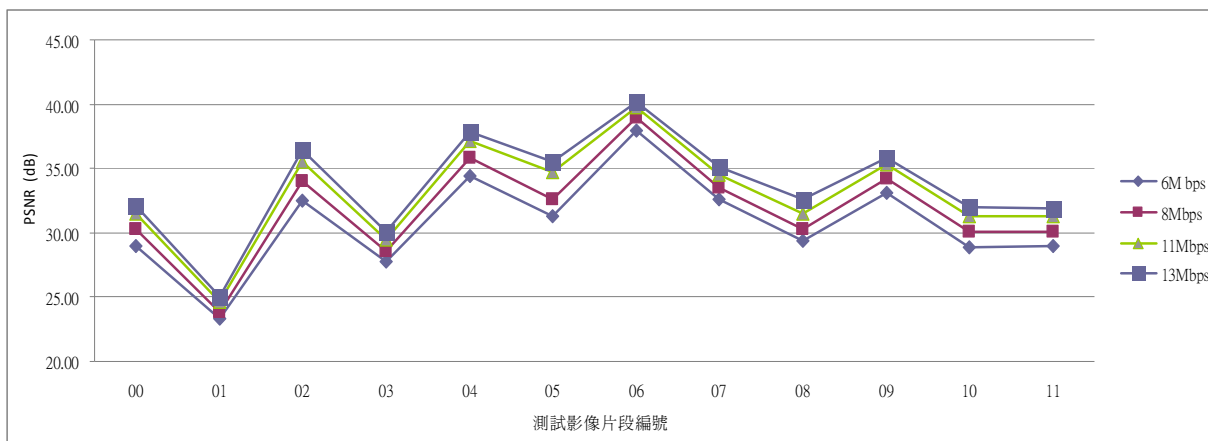
5.4.2. 客觀評量測試系統



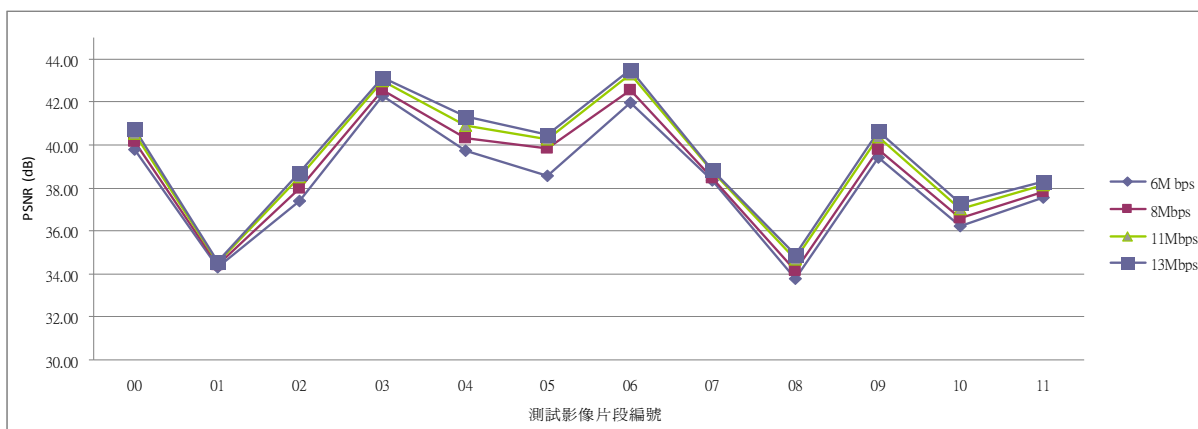
圖表 18、客觀評量測試系統示意圖

5.4.3. 各測試片段 PSNR 數據分析

5.4.3.1. 我們以同於主觀測試用的 12 段測試影像片段進行測試，並分成 4 組壓縮率來進行畫面亮度(Y)及彩色(C) PSNR 的量測，得到的平均數據統計如圖表 19 及圖表 20：



圖表 19、測試片段在不同壓縮率的 PSNR(Y)值



圖表 20、測試片段在不同壓縮率的 PSNR(C)值

5.4.3.2. 以上統計數據由 Omnitek PQA 進行各測試片段量測，不同 H.264 壓縮率量測所得 PSNR 的平均值初步分析發現，雖然不同片段有不同的 PSNR 值，但 h.264 壓縮率提高時，也會使 PSNR 升高，所以 PSNR 與壓縮率之間是成正比的關係。

5.4.3.3. 但是由片段 01 卻可以發現，這一段水波即使提高了壓縮資料量，卻無法有效的提高 PSNR 值，可以驗證有些內容即使在 H.264 高資料量壓縮下仍無法有效的保留住畫面真實度。

5.4.3.4. 而片段 08 在亮度 PSNR(Y)值還算正常，但是在色彩 PSNR(C)值卻相當的低，研判可能是棚內節目在暗處的雜訊太高，現場燈光光源不均勻造成，所以即使畫面並不複雜，但是壓縮後卻造成一塊塊彩斑等等的雜訊。

5.4.3.5. 依據相關研究文獻⁵說明，有關主觀測試所取得的 MOS 值與客觀測試的 PSNR 值有一個約略的比較表如表格 10，而由該比較表中我們可以更清

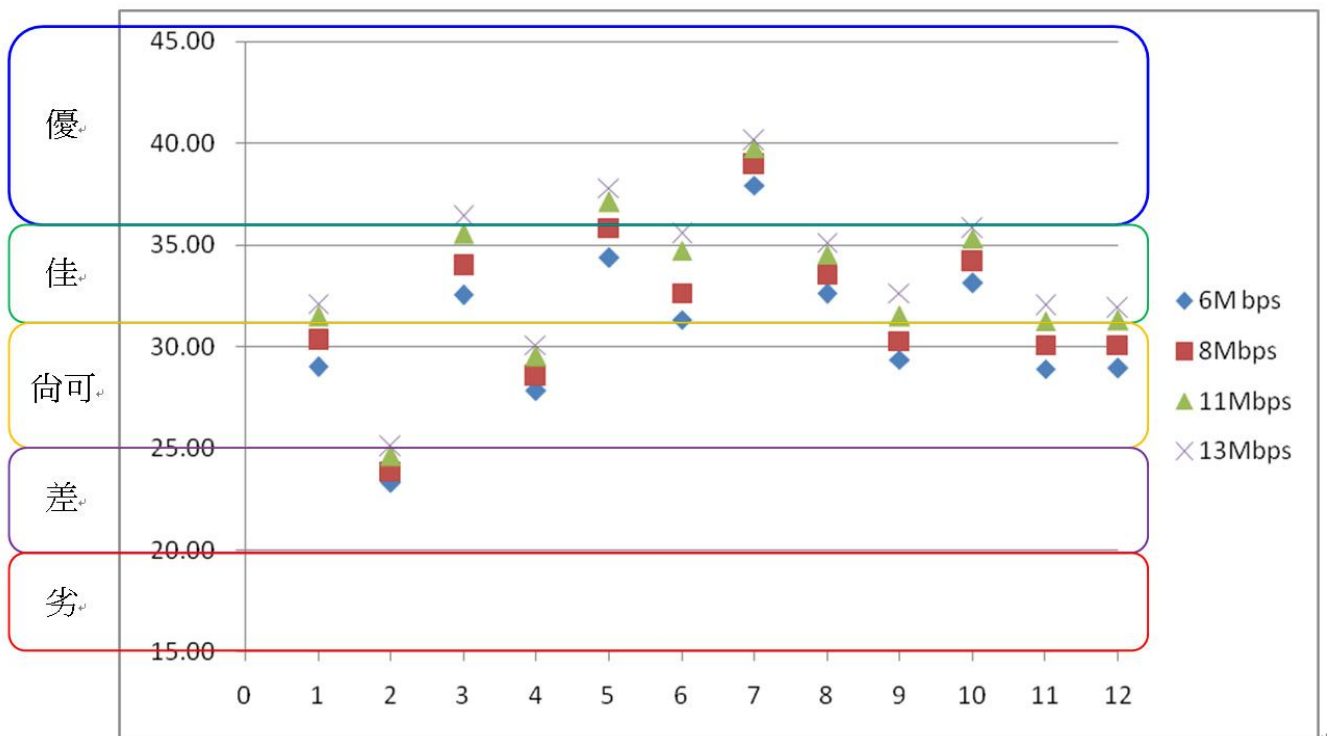
⁵ A Framework for Video Transmission and Quality Evaluation, P.8 - Jirka Klaue,

楚的去比對到這兩種不同的測試畫質的評量方法所表現的意義。而利用這個對應的數據值範圍，將 PQA 量測所得的數值繪製成圖表 21 做為比對分析。

表格 10、客觀 PSNR 值與主觀評量 MOS 對應表

PSNR (dB)	主觀評量方法 MOS	
> 37	5	優 (Excellent)
31 - 36.9	4	佳 (Good)
26 - 30.9	3	尚可 (Fair)
21 - 25.9	2	差 (Poor)
<20.9	1	劣 (Bad)

5.4.3.6. 在圖表 21 其中發現 12 組測試片段在 13Mbps 壓縮率的 PSNR 大部分(10 組)落在相對於主觀評量(MOS)「佳」以上，整體來說是一定比較好建議值；而相對的 11Mbps 就有將近四組在「佳」與「尚可」之間，所以在 11Mbps 以下壓縮率的畫質表現以 PQA 量測到的 PSNR 也較不理想。



圖表 21、PSNR 量測結果跟 MOS 對應表比對圖

5.5. 小結：

- (1). 我們所量測每一個測試影像片段所得的 PSNR，若要大部分都落在相對於 MOS 的「佳」以上時，所要選擇的壓縮資料量最少要 11M bps 以上，最好是 13M bps 會更好。

- (2). 測試片段 01 的 PSNR 量測值即使用高資料量(13M bps)處理後相對其他片段仍要低得很多，可見片段 01 水波漣漪的內容在 H.264 壓縮處理時將會是一大挑戰，似乎是 H.264 的壓縮技術的弱點，但是相對於人眼的主觀評量時，卻似乎是可以被忽略的?這也是主觀與客觀評量之間的差異。
- (3). 在這次主觀評量與客觀評量各自平均所得的結果，一如前面所提及的，兩個結果並不盡完全相同，畢竟人眼的特性與評量時專注重點與 PQA 設備以點對點的比對等特性是有所不同，所得結果因而一定有所差別。但是即便如此，這其間也還可以整合出兩個評量結果共同的特點，得到一些在影片壓縮傳輸的資料量上實質的建議。
- 每個不同內容的影片所需的 H.264 壓縮資料量不同，也就是一個固定的資料傳輸流量是不能滿足不同 HDTV 內容需求，但現實條件下，壓縮資料量不可能沒有限制，必須要跟畫質之間必須取得相當的妥協。
 - PSNR 值會隨著 H.264 壓縮資料量的提高而變高，對客觀評量而言，資料量對 PSNR 而言是成正比，而在主觀評量上壓縮資料量通常也是使得畫質好壞的關鍵之一，但卻不一定是絕對。

陸、結論

6.1. 對我 HiHD 頻道壓縮傳輸資料量的建議

由本次的高畫質影像的評量測試，不論是耗費人力時間的主觀評量方法或是借助以設備儀表進行的客觀評量方法，我們基本上已經得到一個令人滿意的測試結果。眾所皆知，畫面品質跟壓縮的傳輸資料量有一個相對關係，但並非是一個絕對，如何取得一個最大效益是必須要不斷去挑戰它，所以我們宏觀整個 HiHD 高畫質電視測試頻道的相關條件與相關要素，可以整理出一個建議值供參考如表格 11。

如果高畫質電視測試頻道的總承載量(Payload)是規劃為 14.9Mbps 的大前提下，則需看整個高畫質測試頻道上如何規劃營運模式，如果僅規劃一高畫質頻道服務(HiHD Channel)，其節目內容經 H.264 壓縮後也僅為 CBR 形式 (Constant Bit Rate)，所以影像的壓縮資料量建議為 13M bps。

若是在 6Mhz 頻寬內計畫傳送一 HDTV 頻道加上一 SDTV 頻道時，這兩組頻道皆以 H.264 VBR(Variable Bit Rate)進行資料壓縮，並以「統計多工」計算後更有效彈性調配壓縮效率時，則建議該 HDTV 影像的壓縮資料量建議為 10~12M bps，SDTV 為 1~3Mbps。

表格 11、HiHD 頻道頻寬建議參考值

全島 SFN 網路調變條件= 64QAM,CR:2/3,GI:1/4, Payload = 14.9M bps			
方案一	PSI/SI table		500K bps
	1 Ch of HDTV service (with H264 CBR coding)	Video (1080i)	13M bps CBR (H.264)
		Audio (He-AAC)	2 of Stereo ch : 64~96K bps for each pair 5.1 ch: 128K bps
	Data (EPG,Subtitling,SSU...)		Approx. 900k bps
方案二	PSI/SI table		500K bps
	1 Ch of HDTV service + 1 Ch of SDTV service (with H.264 VBR coding & Statistical multiplexing)	Video (1080i)	10~12M bps VBR (H.264 coding)
		Audio (He-AAC)	2 of Stereo ch : 64~96K bps for each pair 5.1 ch: 128K bps
	Video (480i)	1~3M bps VBR (H.264 coding)	
	Audio (He-AAC)	2 of Stereo ch : 64~96K bps for each pair	
Data (EPG,Subtitling,SSU...)		Approx. 900k	

6.2. 提高節目壓縮資料量，影像品質真能提高多少？

自從 HiHD 開始於北高地區進行試播以來，在 HiHD 頻道的節目壓縮資料量就常受到討論、有些人士會以此來揣測頻道上所表現的畫面品質，也常單單以資料傳輸量就直接跟日本 BS-hi 頻道來評論畫質表現，整體而言，高畫質電視節目在進入廣播頭端系統進行的壓縮程序後，由於必須要大量降低原有節目資料量，由原本的 1.5G bps 降低到 10~13Mbps 資料量，所以都必須經過破壞性壓縮資料處理的方式，這確實會影響到原始畫面的品質，但是因為壓縮技術不斷更新及提升效率，以目前 H.264 的壓縮技術及因應針對人眼特性而設計，與原始未壓縮的節目內容比對結果也可以盡量的達到無法分辨的程度，我們在這一一次的評量測試上，也再一次得到這方面驗證的結果。

然而是什麼決定壓縮影像再解碼後的最高可達到的品質？是傳輸資料量嗎？答案是否定的。壓縮影像後解碼的最高可達到的品質應由壓縮編碼之前的影像品質就決定了第一步，或者說由原始信號來源的品質決定第一步。在原始影像內容→壓縮編碼→傳輸→解碼接收的廣播傳輸路徑中，後 3 個步驟過程都是具有破壞性而有損畫質的處理，或者說即便是最好的解決方案，基本上也僅能維持原來影像的品質而已。不管壓縮編碼能力是多麼的優秀，所用的傳輸資料量有多大，經壓縮再解碼後的影像的品質絕對不可能超過原始影像品質。在接收端，不管採用什麼方法，是降低雜訊、採用倍頻掃描或採用其他任何加強影像品質方法也好，顯示的影像解析能力都不可能超過傳輸過來的信號的最大解析能力。

正確的說，僅僅靠數位電視傳輸及壓縮/解壓縮的這一部分是解決不了所謂畫質的問題，要構成完整的高畫質電視製播環境、人員專業素養提昇及整體系統的品質管制來實現。數位電視是整體品質系統的要求，不單只是信號壓縮品質的概念。製播流程的措施跟環節都必須有很好的品質控管。如此也才能將一個高畫質的節目內容真實呈現在收視戶的眼前。

6.3. 國外相關研究分析對照及文獻說明

全世界各國最近這幾年來在 HDTV 的發展上也漸漸加速，由其在歐洲的 DVB-T 數位電視網路漸漸建設完成時，HDTV 的節目內容服務也有越來越多趨勢，然而原本一個 HDTV 服務在由原本 MPEG2 改用 MPEG4/H.264 壓縮格式下，大幅的提高了原傳輸通道頻寬的利用率，也使得各國的 HDTV 不再是那麼個耗費資源的一個服務。像法國地面波數位電視廣播服務中，共計一個 HD Multiplexer 有 4 個 HDTV 服務頻道進行統計多工作業(Statistical MUX)，平均一個 HD 頻道的資料量約為 7.3M bps。

歐洲 EBU 組織 2009 年 2 月提出 TECH 3334「Accommodation of HDTV in the GE06 Plan」報告中對 HDTV 節目採用 MPEG/H.264 壓縮格式而提出的建議壓縮資料量，對於一個 MUX 可整合很多頻道 1080i 的服務內容而可以進行統計多工作業時，一個 HD 頻道建議是 8-10M bps 的壓縮率，若是一個 MUX 裡只有單一個 HD 頻道時，最低的建議是 12.8M bps 的壓縮率(如圖表 22)。

日本國內總務省網站上公告一份 2006 年研究資料顯示，同樣在 DVB-S2 中將大量運用 H.264 壓縮技術來傳送高畫質影音節目，他們亦依據 ITU-R BT.500 主觀

評量方式得到了一些評量的結論，不約而同的提出了 HDTV 傳輸建議值為 13Mbps 可滿足現況需求，如圖表 23。

最後，當我們在計算一個HDTV服務所需的傳輸資料量時，除了影像縮需要的傳輸率之外，另外還要加上聲音用的資料量(大約是預留500K bps給5.1音軌使用)和電子節目表(EPG)、互動資訊服務(iTV)、隱藏式字幕(Subtitle) 等等資訊內容廣播用的空間。

HDTV Service... Roll out considerations – few more facts...

- DVB-S2 provides up to **35%** gain over DVB-S
- H.264/AVC (MPEG4) was proven to provide **~50%** coding gain over MPEG-2 for delivery rates.
- 720p/50 provides **20%** benefits in distribution than 1080i/25
 - **EBU recommendation R124**
- Minimum (video) bit rate to provide HD quality (from EBU tests – **EBU BPN085-087**):
 - 1280x720p/50 – **10 Mbps**
 - 1440x1080i/25 – **12.1 Mbps**
 - 1920x1080i/25 – **12.8 Mbps**

(Measured over several sequences and using 5 different encoders)
- Using statistical multiplexing helps balance the rate on other channels (more HD services or higher quality at lower costs)

EBU TECHNICAL - your reference in media technology and innovation

(資料來源：HDTV services *Trends and Implementations* (2009) / Adi Kouadio/EBU)

圖表 22、歐洲EBU提出的最低資料量需求

2. 映像符号化部 (H.264) の技術的条件

主観評価の結果、HDTV:13MbpsでITU-Rが定める放送画質への要求条件に概ね適合することを確認

検証結果

フルスペックのHDTV(以下「フルHDTV」)を対象として、ITU-R勧告に基づく画質評価方法に則り、ITU-Rが定める放送画質への要求条件^(注)に適合できるビットレートを確認。
 ⇒ 現時点では、13Mbpsにて概ね適合する画質が得られることを確認。
※ 13Mbps未満のビットレートにおいて、将来、ITU-Rの要求条件に適合する可能性を否定するものではない。

ITU-Rの要求条件への適合性

ビットレート	評価画像の3/4以上で原画からDSCQS 12%以内の劣化	全ての評価画像で原画からDSCQS 30%以内の劣化
7Mbps	—	◎
9Mbps	—	◎
13Mbps	○	◎
17Mbps	◎	◎

- ◎ 要求条件に対し余裕をもって適合する
- 要求条件に対し概ね適合する
- 現状では要求条件に適合しないが、今後の画質改善が期待される

パラメータの設定

項目	仕様
プロファイル	Highプロファイル (Mainプロファイルを含む)
レベル	映像入力フォーマット: 480i、480p、720p、1080iにそれぞれ対応する3、3.1、3.2、4を採用
映像信号 (カラーマトリクス等)	地上デジタル放送、BSデジタル放送及び広帯域CSデジタル放送の規定(ITU-R BT.709準拠)を準用

(注)ITU-R 勧告 BT.1210等から選んだ少なくとも4種類の評価画像のうち、75%の評価画像にてDSCQSスケールで品質差12%以内であること。
 残りの評価画像についても30%以内の品質差であること。
※DSCQS = Double-Stimulus Continuous Quality Scale (二重刺激連続品質尺度):ITU-R勧告BT.500-11で定める主観評価法


(資料來源:

http://www.soumu.go.jp/menu_news/s-news/2006/pdf/060720_9_1.pdf)

圖表 23、日本總務省網站在 HDTV 傳輸資料量中的建議值資料

6.4. 未來技術演進，效率與品質會更好

由於晶片的運算能力越來越強大，所以對於 H.264 壓縮的效能也日益提升，有些製造商為了追求更高的壓縮效能，更自行開發大型 ASIC 專用晶片做為壓縮器的運算核心(例如 Thomson 的 Mustang Chipset)，有些廠商也提出以 FPGA 來作為壓縮的硬體架構(如圖表 24)，在不斷開發新技術時，只要修正時將軟體升級即可。這樣機制也才能讓營運商能藉由軟體升級的方式將壓縮的效率不斷提昇，而不用一直汰換壓縮設備來提升效能所造成的成本損失。



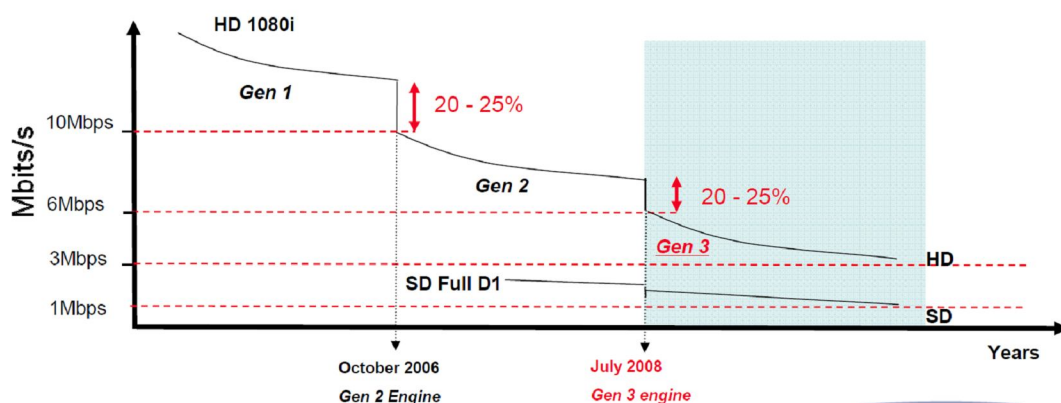
- **Powerful FPGA hardware, allowing combination of Processor-Hungry Algorithms and tools.**
 - Power equal to 60 1Gz DSP, or 10 Quad-core Intel 12GHz processor.

(資料來源: Ateame 公司/ ATEME Contribution Solution 2009 presentation)

圖表 24、ATEME 公司所提出壓縮器內 FPGA 硬體效能

據了解，目前全世界新建的高畫質電視頻道，不論是經由地面波、衛星或是網路的傳輸路徑，絕大部分都是採用 H.264 壓縮格式，相信影像經 H.264 壓縮的數位資料量也在很多人的努力下效能會一再提高，有些廠商的目標是希望下一代產品可以將資料量再降低 20%~25%(圖表 25)，而影像品質也會達到令人滿意

的程度。



(資料來源：Ateme 公司/ ATEME Contribution Solution 2009 presentation)

圖表 25、製造商希望下一代 H.264 壓縮器的產品效能目標

目前 H.264 壓縮除了用於廣播(Distribution/Broadcast)之外，目前更積極推向中繼傳輸(Contribution)用的格式，其實在 2004 年 JVT 工作組就對 H.264 做了即為重要的補充擴展，通過擴展部分使 H.264 標準得以涵蓋高傳真視訊壓縮的應用領域，這個擴展部分被簡稱為 FRExt (Fidelity Range Extensions)。FRExt 部分的應用能夠使 H.264 支援專業級的高品質影像應用和高解析度的視訊壓縮等，同時標準本身在原有的 Baseline、Main、Extended 三個類的基礎上增加了四個新的類：High (HP)、High 10 (Hi10P)、High 4:2:2 (H422P)、High 4:4:4 (H444P)。FRExt 相對於早期 H.264 壓縮的改善主要如下：

- 引入 8x8 的整數變換、量化矩陣、8x8 的幀內預測等壓縮技術；
- 影像取樣量化值超過 8 bits，最高可達到 12 bits，達到了數位電影的要求；
- 增加了 4:2:2 與 4:4:4 的取樣格式，其中在 H444P 中還能夠支援高效能的無損失區域壓縮編碼，並且在壓縮 RGB 影片影像時採用了整數色彩空間轉換方式，避免了原先色彩空間變換時的舍入誤差；

支援更高的位元速率和影像尺寸 (4K × 2K)，詳見表格 12 和表格 13。

表格 12、H.264 的分級特性

級(Level)	最大支援影像格式	非 FRExt 最大資料量 (Mbps)
4	1920x1080	20
4.1	1920x1080	50
4.2	1920x1080	50
5	2Kx1K	135
5.1	4Kx2K	240

如果採用 FRExt Profile 編碼，則表格 12 中的最大資料率應用表格 13 中的係數進行相乘。

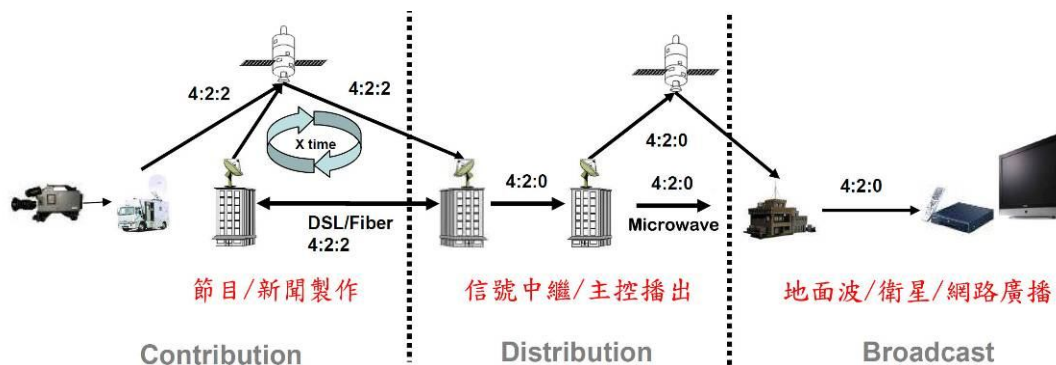
表格 13、係數

FRExt Profile	相乘係數
HP	1.25
High 10 (Hi10P)	3
High 4:2:2 (H422P)	4
High 4:4:4 (H444P)	4

- 例如非 FRExt Profile時等級4的最大資料率為20Mbps，則在 FRExt Profile 中的 High444時最大資料率可為80Mbps。

由於HP能夠比MP (Main Profile) 更為有效地提高編碼效率，並且不會過多的增加實現複雜度，因此預計HP將很快全面取代MP (Main Profile) 的地位。

H.264 在壓縮效率上已經大大超過了以往的影像壓縮標準，再經過 FRExt 的改善，更能夠廣泛地支援大尺寸、高品質的影像壓縮應用。現今，高品質、大尺寸的影像應用的需求迫切，數位高畫質電視壓縮技術在很多的領域得到了廣泛的關注和應用。H.264 以追求影像壓縮效率為目標，通過複雜的壓縮、編碼演算法，利用有限的頻寬傳輸盡可能追求高品質的影像，主要適用廣播電視和多媒體等領域，值得注意的是當 FRExt 補充入 H.264 之後，它也適合於中繼傳輸用高傳真的視訊壓縮應用。



圖表 26、H.264 壓縮技術應用於整個製播環境示意圖

柒、 参考文献

1. T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 560–576, July 2003.
2. S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 645–656, July 2003.
3. T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 657–673, July 2003.
4. T. Wedi, "Motion compensation in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 577–586, July 2003.
5. M. Flierl and B. Girod, "Generalized B pictures and the draft JVT/H.264 video compression standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 587–597, July 2003.
6. T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 688–703, July 2003.
7. Schafer Ralf, Wiegand Thomas, Schwarz Heiko. "The emerging H.264/AVC Standard EBU Technical Review", Jan.2003
8. M. Karczewicz and R. Kurçeren, "The SP and SI frames design for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 637–644, July 2003.
9. H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-Complexity transform and quantization in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 598–603, July 2003.
10. P. List, A. Joch, J. Lainema, G. Bjøntegaard, and M. Karczewicz, "Adaptive deblocking filter," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 614–619, July 2003.
11. J. Ribas-Corbera, P. A. Chou, and S. Regunathan, "A generalized hypothetical reference decoder for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 674–687, July 2003.
12. Mathias Wien, "Variable Block-Size Transforms for H.264/AVC" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 604–613, July 2003.
13. Michael Horowitz, Anthony Joch, Faouzi Kossentini, "H.264/AVC Baseline Profile Decoder Complexity Analysis" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, pp. 704–716, July 2003.

14. T. Wiegand and G. J. Sullivan, "Draft ITU-T Recommendation H.264 and Final Draft International Standard of Joint Video Specification (ITU-T Recommendation H.264 | ISO/IEC 14496-10 AVC)", Joint Video Team of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q.6 Doc. JVT-G050, Pattaya, Thailand, Mar. 2003.
15. JVT Document JVT-C028
G. Bjontegarrd and K. Lillevold, "Context -adaptive VLC Coding of Coefficients", Fairfax, VA, MAY 2002
16. Yao Wang, Jorn Ostermann, Ya-qin Zhang. 視頻處理與通信. 北京: 電子工業出版社. 2003。
17. X. L. Ce Zhu, and Lap-Pui Chau. Hexagon-Based Search Pattern for Fast Block Motion Estimation, IEEE Trans. on CSVT, vol. Vol.12, 2002: 349-355.
18. Ghanbari, M. The cross-search algorithm for motion estimation, Communications, IEEE Transactions on, Volume: 38 Issue:7, July 1990: 950-953.
19. Chang, Y.-C., D. G. Messerschmitt, T. Carney, and S. A. Klein. Delay cognizant video coding: architecture, applications, and quality evaluations. Forthcoming in IEEE Trans. Image Processing.
20. Iain E. G. Richardson. H.264 and MPEG-4 Video Compression. Aberdeen, UK. 2003.
21. Joint Model for Non-normative Aspects of Advanced Video Coding. Study of ISO/IEC 14496-6 / PDAM6: 2003 (E).
22. H.264 運動估值與網路視頻傳輸關鍵問題研究, 清華大學, 畢業論文, 2002.
23. 數位影像通信原理與技術. 北京: 國防工業出版社, 2000: 89-92
24. 丁貴廣, 計文平, 郭寶龍, Visual C++6.0 數位影像編碼, 機械工業出版社, 2004。
25. 朱秀昌, 劉峰, 胡棟, 數位影像處理與影像通信, 北京郵電大學出版社, 2002。
26. Marta Karczewicz The. SP- and SI- Frames Design for H.264/AVC. IEEE Trans. Circuits and Systems for Video Technology, Vol 13, No.7, July 2003
27. TML 8.7 Software <ftp://standard.pictel.com/video/h261/>
28. 萬萍, 陳仁雷, 王海嬰, H.264/AVC 中的 SP/SI 幀技術, 電視技術, 2004.1, p22-25
29. 畢厚傑主編, 新一代視訊壓縮編碼標準-H.264 2004 p.97-164
30. ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures, 2002
31. GY/T 134-1998 數位電視影像品質主觀評價方法

32. 黃文輝 陳仁雷 張家謀，數位視訊影像品質客觀測量方法的改進與實現，
2006.11.24
33. A Framework for Video Transmission and Quality Evaluation P.8 - Jirka
Klaue, Berthold Rathke, and Adam Wolisz

附錄 LG LCD TV 42LG60 規格

Model Number	42LG60
Flat Panel Type	LCD TV
LCD Response Time	5 ms
Screen Size	42 inch
Contrast Ratio	50,000:1
Display Resolution	1920 x 1080 pixels
Brightness	500 cd/m ²
Output Color	10 million colors
Viewing Angle	178
Adjustable Color Temperature	With Adjustable Color Temperature

