

# 檔案管理局

電子檔案保存管理機制委託服務案

電子檔案轉置技術指引

(1011130\_01)

英福達科技股份有限公司 謹呈

民國 101 年 11 月

# 目錄

壹、	簡介.....	- 1 -
貳、	轉置分類.....	- 2 -
一、	媒體轉置.....	- 2 -
二、	格式轉置.....	- 3 -
參、	轉置品質.....	- 6 -
一、	影像品質評估.....	- 6 -
二、	視訊品質評估.....	- 10 -
肆、	轉置作業程序.....	- 13 -
一、	確認轉置對象.....	- 13 -
二、	進行電子檔案資料備份.....	- 13 -
三、	進行電子檔案安全檢測.....	- 13 -
四、	進行電子檔案格式辨識與檢測.....	- 13 -
五、	確認轉置格式及儲存媒體.....	- 13 -
六、	品質驗證.....	- 14 -
七、	保留詮釋資料.....	- 14 -
八、	資料清除.....	- 14 -

## 壹、簡介

近年來由於資訊系統的普及，電子檔案的數量急遽增加，但由於資訊技術日新月異，電子檔案的儲存格式與儲存媒體幾乎是以 3 至 5 年為一個世代快速轉變，有鑑於此，如何確保被保存的電子檔案能持續地被存取與利用，已成為一個非常重要的課題。目前已有許多長期保存方法與策略被廣泛地探討，但仍以系統保存、模擬與轉置等為主要及常見之保存策略，其中又以轉置最廣為接受與使用。然而電子檔案每經過一次轉置都有可能造成資訊的流失，甚至可能是極為珍貴的資料。

本指引範圍包括電子檔案格式轉置、媒體轉置與品質驗證，說明電子檔案轉置技術，說明轉置分類、時機與相關注意事項。

## 貳、轉置分類

### 一、媒體轉置

所有的電子檔案都必須經由支援該電子檔案儲存格式的軟體解讀呈現並藉由媒體(載體)儲存，如果媒體因故無法取得或讀取，儲存於媒體內之電子檔案也將無法存取與利用，因此，儲存媒體選擇在電子檔案長期保存實務上相當重要，故在選擇儲存媒體時，應該針對各種儲存媒體的使用年限、保存環境、硬體支援程度、普及性、儲存空間及需求等條件妥善評估。

#### (一)媒體轉置時機

1. 媒體本身已屆或將屆保存期限。
2. 儲存設備更新。
3. 原媒體儲存空間不足。
4. 媒體因外力或其他因素面臨損毀之風險，例如：光碟片因受潮嚴重所產生之風險。
5. 因為管理方式改變或應用系統修改或更新而必須進行資料結構調整，如資料夾異動或合併等。
6. 維護成本考量，如舊儲存媒體容量較小體積較大，雖未屆保存期限，但為求維護方便，必須進行轉置。

## (二)媒體轉置注意事項

1. 媒體轉置應建立標準作業程序，產出標準作業文件，  
以備後續查核與作業改善之依據。
2. 媒體轉置前應先進行檔案比對與總和檢查。
3. 媒體轉置必須保持原媒體資料夾結構，如無法保持原  
資料夾結構時，必須詳載調整方式，以提供後續查核  
使用。
4. 媒體轉置後之新儲存媒體應即刻依標準保管方式進  
行統一保管，並注意保存環境不可過於潮濕並應保持  
穩定之溫度。
5. 媒體轉置完成，應檢視總和檢查結果、檔案雜湊值、  
轉置檔案數量及轉置時間等資料，以驗證轉置品質，  
並記載於詮釋資料中。

## 二、格式轉置

電子檔案是以特定格式描述之資料，必須經由應用程式方能予以呈現，例如：Microsoft Office 中的 Word 檔案，其中記載了各段落之文字格式，必須依賴能解讀 Word 格式的應用程式才能將其呈現與輸出；各種應用程式也必須執行在作業系統之上，某些作業系統更只能執行在特定硬體架構上。

### (一)格式轉置之時機

1. 儲存的電子檔案格式為非開放格式。
2. 資訊技術快速發展，電子檔案格式可能過時，或因普及率不高，應用程式廠商可能停止提供支援或服務。
3. 支援電子檔案之應用程式相依之作業系統或硬體架構無法繼續獲得。
4. 由於長期保存之需要，必須將特定格式之檔案轉置為適合長期保存之格式。
5. 配合法令規定。

### (二)格式轉置注意事項

1. 進行格式轉置前，應評估相關應用系統是否需要一併調整或更新。
2. 應建立格式轉置標準作業程序，並產出轉置作業相關文件，以利日後查核及作業改善之依據。
3. 轉置後之目的格式必須符合長期保存要求。
4. 轉置品質檢驗。
5. 應將轉置工具名稱、轉置工具版本、轉置參數、品質驗證結果及歷史版本資訊等轉置相關資訊記載於詮釋資料中，與轉置後檔案一併保存。

6. 格式轉置前，應評估相關檔案是否完整，並應進行格式辨識及掃毒程序。

7. 格式轉置格式限制

- (1) WDL 格式轉置為 PDF/A 格式。
- (2) DOC 格式轉置為 PDF/A 格式。
- (3) TIFF 格式轉置為 JPEG 格式。
- (4) JPEG 格式轉置為 TIFF 格式。
- (5) WMV 格式轉置為 MPEG-2 格式。
- (6) MPEG-2 格式轉置為 H.264 格式。
- (7) DOCX 格式轉置為 PDF/A 格式。
- (8) PS 格式轉置為 PDF/A 格式。
- (9) XLS 格式轉置為 PDF/A 格式。
- (10) EML 格式轉置為 PDF/A 格式。
- (11) PPT 格式轉置為 PDF/A 格式。
- (12) MP3 格式轉置為 WAV 格式。
- (13) GIF 格式轉置為 PNG 格式。
- (14) DOC 格式轉置為 ODT 格式。
- (15) TIFF 格式轉置為 PNG 格式。
- (16) TIFF 格式轉置 PDF/A 格式。

## 參、轉置品質

無論是任何形式之轉置作業，都存在資訊流失的風險，因此必須建立一個評量機制，以確保轉置品質。評量媒體轉置品質可藉由以下幾個方式進行：

- (一) 原始儲存媒體與轉置後儲存媒體檔案總數比對。
- (二) 原始儲存媒體(電子檔案)與轉置後儲存媒體檔案雜湊值比對。
- (三) 是否產出標準作業文件。

評量格式轉置品質的方式可分為客觀(Objective)與主觀(Subjective)兩種準則：

- (一) 客觀準則：以計算轉置前後的量化指標來評估影像品質。
- (二) 主觀準則：以人類肉眼的觀察來比較轉置前後影像品質。

評量格式轉置品質的類型可分為影像品質評估、語言品質評估及視訊品質評估，以下為針對影像品質評估及視訊品質評估類型做詳細介紹。

### 一、影像品質評估

常見的客觀評量方法有影像方面的均方差(Mean Square Error, MSE)、峰值信號雜訊比(Peak Signal to Noise Ratio,



PSNR)、結構相似性指標(Structure Similarity, SSIM)、通用性影像品質指標(Universal Quality Index, UQI)及基於人類視覺系統之峰值信號雜訊比(HVS- Peak Signal to Noise Ratio, HVS-PSNR)的評估方式，其中最廣泛被使用的客觀評量方法是峰值信號雜訊比(Peak Signal to Noise Ratio, PSNR)。

#### (一) 均方差(Mean Square Error, MSE)

均方差為一般影像評估常用的方式之一，求出來的 MSE 值越小，表示輸入影像與複製影像之間的差異少，也就代表品質較好。均方差主要計算方式是將輸入影像以及複製影像的每一個像素點相減，再將其差值的平方加總起來取平均值，取得 MSE 值，其計算公式如方程式下所示。

$$MSE = \frac{1}{MN} \sum_{i=1}^m \sum_{j=1}^n [f(i,j) - \hat{f}(i,j)]^2$$

#### (二) 峰值信號雜訊比(Peak Signal to Noise Ratio, PSNR)

PSNR 也就是峰值訊噪比，經常用作圖像壓縮等領域中信號重建質量的測量方法，它是利用影像信號的最大值與影像中雜訊的比值作為評估的標準，其計算公式如下方程式，其中  $(2^n - 1)^2$  代表是表示圖像點顏色的最大數值，如果每個像素點用 8 位表示，那麼就是 255。在利用 PSNR 算出來的比值越大，代表複製影像與輸入影像越接近，影像品質良好。

$$\text{PSNR}=10\log\frac{(2^n-1)^2}{\text{MSE}}$$

### (三) 結構相似性指標(Structure Similarity, SSIM)

結構相似性指標是用於測量在兩張影像之間相似性的一種方法。SSIM 指標係改進以前提出的通用性影像品質指標 (UQI) 模型，而可以被視為一個完美的影像品質量測。

此一指標模型可以透過比較誤差敏感性理論 (Error Sensitivity Philosophy) 而得到瞭解。首先，誤差敏感性接近估計察覺錯誤確定影像降低的數量，考慮影像降低而察覺在架構性訊息變化的變化。其次，錯誤差敏感性理論由下而上的方式，相關功能類似早期 HVS 系統。而此模型是由上至下，模仿 HVS 的功能性，避開在以前部分提及臨界值的問題，因為它不倚賴臨界值確定察覺的變形數量。並降低相互作用問題，因此結構相似性指標主要在評估兩張影像間的結構變化。

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

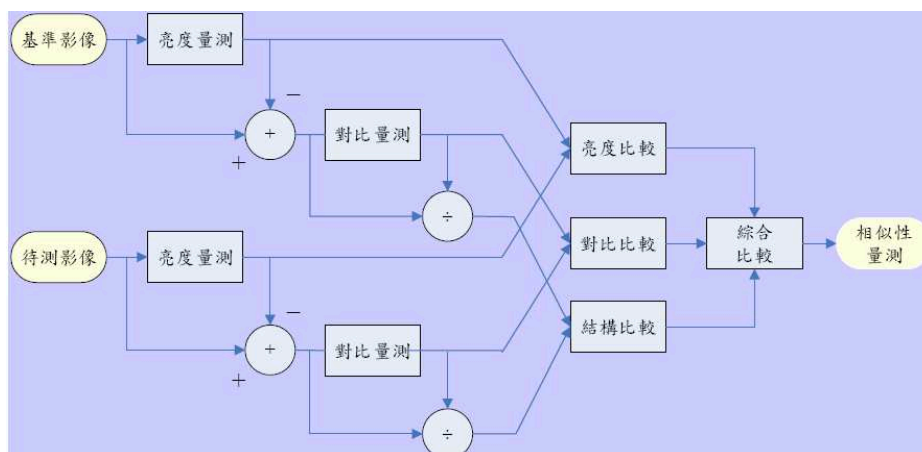


圖 1 結構相似性量測系統圖

傳統客觀品質指標是以統計的方法計算出有關於整張影像整體灰度值誤差的總和，而結構相似性指標不單僅是利用計算灰度值誤差的方法，並以模擬人眼視覺系統(HVS)，來進行影像品質評估結果，而由結構相似性指標計算結果卻可得到不同的指標值，並且得到符合人眼視覺的評估結果。

#### (四) 通用性影像品質指標 (Universal Quality Index, UQI)

通用性影像品質指標，對於各類型影像處理均可適用，設計上是以任何影像失真模型均含有三種因素：相關係數降低、亮度改變及對比改變。

模型評估影像的最佳及最差的量化值均介於-1 至+1 之間，因此可得到比較具體的評估值。

$$Q = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \cdot \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \cdot \frac{2\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

(五) 基於人類視覺系統之峰值信號雜訊比 (HVS- Peak  
Signal to Noise Ratio, HVS-PSNR)

HVS-PSNR 不同於傳統的 PSNR，因為 HVS-PSNR 加入了影像評估的前置處理，在計算每一個像素點的誤差值後，將其通過高斯低通濾波器，模擬人類視覺的低通特性，再計算影像失真的幅度，可以得到接近人眼所察覺到的影像評估值。HVS-PSNR 及 PSNR 一樣，評估值越大，代表轉置影像的影像品質越好。

## 二、視訊品質評估

視訊品質的評估方法可以分為 3 類，分別為：Full-Reference、Reduced - Reference 及 No-Reference，以下為 3 類評估方法的介紹。

(一) Full-Reference

一種參考原始與經處理過後視訊內容的視訊品質評估方法。這類的方法是假設在測量畫面品質的時候，同時也有原始視訊處理前的畫面可供參考。在這類方法中，主要偵測在時間上因 DCT 量化的變化結果造成的編碼誤差，同

時也考慮的時間軸上的後遮蔽現象。此外，許多方法延伸了 SSIM 的做法，在其誤差模型中除了計算每張獨立畫面區塊中的平均值與變異數之外，同時也針對每個信號分別計算其 SSIM 的結果，之後再利用動量估測為結合的權重產生最後的結果。

## (二) Reduced-Reference

一種參考部分原始與部分經處理過後視訊內容的視訊品質評估方法。有時候視訊品質的比較並沒有辦法完整得到原始視訊內容的資訊，例如在視訊串流系統中，我們不可能傳送完整的原始視訊內容給接收端要求其比較正在觀賞的視訊品質。因此，Reduced-Reference 的方法透過事先抽取部分原始視訊內容的特徵值，再將抽取出來的特徵值傳送給比較端後，比較端也同樣抽取相同的特徵值，再比較兩者特徵值得衰減情形比較視訊的品質的好壞。

## (三) No-Reference

一種只參考經處理過後視訊內容的視訊品質評估方法。大部分的應用環境下，我們幾乎都沒有辦法得知原始視訊內容來比較視訊品質，例如觀看網路視訊的時候，我們不可能還有原始影片可以參考，或者是拿到一部 DVD 視訊時，我們也沒辦法得知其原始視訊內容，所以 No-Reference 利用結果影片中的誤差結果去推估視訊的品

質好壞，不需要參考編碼端的原始影片，也因此 No-Reference 的方法最為廣泛應用在大部分的狀況下評估視訊品質。

## **肆、轉置作業程序**

### **一、確認轉置對象**

藉由人工分類或系統自動比對搜尋轉置對象，整理轉置清單並擬定轉置計畫。

### **二、進行電子檔案資料備份**

將欲轉置資料進行複製產生副本，利用檔案個數比對、總和檢查、檔案雜湊值比對等方法檢測轉置資料副本是否與原始檔案相同。

### **三、進行電子檔案安全檢測**

使用病毒掃描軟體與惡意軟體掃描工具針對轉置資料副本進行檢測，以確保系統安全與資料安全。

### **四、進行電子檔案格式辨識與檢測**

針對轉置資料副本檔案使用電子檔案保存工具進行格式辨識與檢測以確定原始檔案格式；如來源檔案為數位內容封裝檔則同時進行封裝檔格式檢測，以檢測封裝檔版本與正確性。

### **五、確認轉置格式及儲存媒體**

參考電子檔案長期保存知識庫系統以選擇適當的轉置格式

及儲存媒體，或利用電子檔案保存工具取得相關說明。

## 六、品質驗證

如為媒體轉置則使用檔案個數比對、總和檢查、檔案雜湊值比對或媒體封裝檢查進行品質驗證，如為電子檔案格式轉置則使用電子檔案保存工具進行品質驗證。

## 七、保留詮釋資料

在轉置的過程中擷取來源檔案屬性、轉置工具、轉置工具版本、轉置時間、轉置檔案個數、驗證說明、原始檔案格式及轉置後檔案格式等相關資料做成詮釋資料並與轉置後資料一併封裝保存。

## 八、資料清除

完成轉置並通過品質驗證後，轉置過程中使用之暫存資料應予以清除，如原媒體已屆保存年限或無法繼續使用，應予以銷毀。