

結合人臉偵測之接縫雕刻法於內容感知影像不等比例縮放之研究

莊靜芬 周佳毅 李孟桓 吳俊霖*

國立中興大學資訊科學與工程學系

jlwu@cs.nchu.edu.tw

摘要

由於科技的快速發展，面對顯示介面的多樣化，我們需要更有彈性的影像尺寸。過去若要不同的影像尺寸，會採取等比例影像縮放、一般非等比例影像縮放，或傳統非等比例影像縮放，後兩者雖然較有彈性，但限制仍然很多，尤以影像重要主體變形、影像內容的破壞最為重大。若要保護影像重要主體，我們需要「內容感知影像尺寸重製」(content-aware image resizing)，此方法原理是利用人類視覺特性，對影像不連續處，也就是邊緣(edge)的部份，較為敏感，於是，我們把目標放在移除不易被注意、與周邊融為一體的那些映像點，也就是影像中視覺上較平緩的部分。

接縫雕刻(seam carving)可以完成內容感知影像尺寸重製，接縫(seam)的挑選取決於映像點的能量值(energy value)；一個映像點的能量值，則取決於該映像點的邊緣多寡、變化程度、平緩程度。然而經反覆實驗發現，若輸入之影像背景複雜或影像中較不重要的部分邊緣(edge)多且與重要主體邊緣數量落差大，其式將會誤判影像重要主體為不重要的部分，致使程式選擇的接縫大多經過重要主體，如此一來不論是移除或者加入接縫，都將造成影像重要部分的變形，成果影像將不合於使用。為了解決此問題，本論文以一般接縫雕刻為基礎，加入人臉偵測後做人臉及人身保護。實驗結果顯示，所提演算法明顯較原先的一般接縫雕刻改善許多。

關鍵詞：接縫雕刻、內容感知、影像尺寸重製、人臉偵測。

1. 前言

科技快速發展，各種面向的需求隨之產生，人們使用的工具因而有了重大變異，影像的顯示介面也愈加多樣。從最早的紙張，到後來的電腦螢幕、相機螢幕，以及近幾年成為新一代通訊主流媒體的智慧型手機，顯示影像的工具已不像從前只有特定幾種，各種顯示介面對影像會有不同的尺寸需求，而同樣的工具也隨著廠牌、系統的不同而有非常多樣的規格。面對顯示介面的多樣化，我們需要更有彈性的影像尺寸；提到影像尺寸的縮放，最基本有兩種方法：等比例縮放及非等比例縮放，其中又以

非等比例縮放較具彈性，但進行非等比例縮放後，影像的內容必定會產生變形，若影像內容包含重要主體物件，這樣的變形將突兀且不可接受，此特性是傳統非等比例縮放在使用上的重大限制。

除了顯示介面，設計與印刷業對影像尺寸彈性的需求也是始終存在的。設計時通常只設計一種尺寸版本，但印刷時則需考慮內容物的大小、形狀，而使用好幾種不同的尺寸，原本設計的影像比例可能也會因而出現變動。然而，若影像的設計或所要印刷的物件較複雜，非等比例縮放的變形可能就會出現突兀的結果，造成印刷時排版的困擾，或甚至需要重新設計，諸多不便。

兼顧提供更有彈性的影像尺寸與保護影像內容物的「內容感知影像尺寸重製」，在大部分的情況下，可以於影像縮放時，保護視覺上較容易注意到、較重要的部分；然而此方法仍有其限制，若我們欲操作的影像背景複雜，很可能誤判影像重要主體為不重要、可變形或可刪除的部分，如此一來此方法亦不完全合於解決非等比例影像縮放之重要主體變形問題。人物影像中，以「內容感知影像尺寸重製」為基礎，若能偵測人臉進而保護整個人身，便能真正在保護主體的情況下進行非等比例影像尺寸重製。

本論文分兩個部分解決影像進行非等比例縮放時所產生的重要主體物件變形問題。首先，由於人類在視覺上通常會注意到影像中邊緣較多的部分，而忽略邊緣較少、較平緩的部分，所以我們希望依據邊緣的多寡，初步判斷一張影像哪裡需要被保護、哪裡則較不重要，可以以接縫(seam)為單位進行刪除，直至我們預期的尺寸。

接著我們使用人臉偵測演算法來改善內容感知影像尺寸重製的缺失、突破其限制，使此方法能為更廣泛的應用。由於背景或影像中較不重要的部分亦可能是邊緣較多的地方，若影像主體為人，產生的變形將難以接受；我們進行人臉偵測與人身保護，如此一來即便是在該種情況下，仍能夠保護人身且做到具有彈性的非等比例影像尺寸重製。綜言之，本論文完成以下三項目的：

- (1)以邊緣數量為依據，初步判斷影像中之重要主體物件。
- (2)特殊影像採用人臉偵測及人身保護，避免人像產生非等比例變形。
- (3)使影像順利進行內容感知影像尺寸重製的非等比例縮放。

2. 相關研究

由於科技的快速發展，人們每天都會面對各式各樣的影像，影像在各行各業、各個方面都扮演重要角色，人們對影像的需求也愈加提升；科技的發展帶來了更方便的工具，影像的顯示裝置也更加多樣，非等比例影像尺寸重製成了重要的、我們所必須使用的影像縮放手法，但傳統的非等比例影像尺寸重製仍有諸多限制，訂定了希望可以提供更有彈性、更有智慧，能夠滿足更多複雜影像的非等比例影像尺寸重製工具這個目標，我們採用接縫雕刻 (Seam Carving) 配合人臉偵測與人身保護，達成更能保護影像重要部分的「內容感知影像尺寸重製」。

2.1 傳統之非等比例影像尺寸重製

我們實現內容感知影像尺寸重製的方法是審慎地移除映像點。然而，如何選擇我們所要移除的映像點呢？直覺上會把目標放在移除不易被注意、與周邊融為一體的那些映像點，也就是影像中視覺上較平緩的部分；相反地，影像中如果出現邊緣 (edge) 較多、變化較多的部分，通常我們將視之為影像中的重要部分，因為在大部分的情況下是這樣的；我們會採取一些措施保護影像中的此類部分。由於邊緣的多寡在這份報告裡是很重要的影響因素，我們需要事先定義能量 (energy)：一個映像點的能量值，取決於該映像點的邊緣多寡、變化程度、平緩程度；這引導出以下這個能量函數，透過對影像做一階微分所得之影像變化率，也就是梯度 (gradient)，我們將此梯度當做影像的能量函數 e ：

$$e_1(I) = \left| \frac{\partial}{\partial x} I \right| + \left| \frac{\partial}{\partial y} I \right| \quad (1)$$

給定一個能量函數，假設我們需要縮小影像的寬度，可以想到數種對策來達成，以下我們介紹移除低能量映像點 (remove the pixels with lowest energy)、每列等數量移除低能量映像點 (remove an equal number of low energy pixels from every row)、自動裁切 (Auto-cropping)、整行性移除低能量映像點 (remove whole columns with the lowest energy) 這四種方法。

(a) 低能量映像點之移除

「低能量映像點之移除」是保存能量的最理想對策，它的操作方法是在上升序列中，移除能量值最低的數個映像點。在這裡的保存能量 (preserve energy) 的意思是留下具有高能量值的映像點。但在此方法中，我們每列可能移除不同數目的映像點，這將會破壞影像原本的矩形形狀；由此可知，使用「低能量映像點之移除」來縮小影像的寬度，成果影像可能不會是我們所習慣的矩形影像，不適合做普遍性的運用。

(b) 每列等數量移除低能量映像點

若想防止前節 (低能量映像點之移除) 所提之影像形狀被破壞、縮小影像之寬度然而保持影像的矩形形狀，可以每列移除相等數目的低能量映像點。但此方法雖保護了影像的矩形形狀，卻可能會產生 Z 字型效果 (zigzag effect)、影像中出現曲折的結果，破壞了影像的內容。若成果影像產生人眼可察覺的 Z 字型效果，雖達到縮小影像之寬度的目的，成果影像仍然會因不自然的 Z 字型效果破壞而不合於使用。所以我們可知，此方法亦不適合做普遍性的運用。

(c) 自動裁切 (Auto-cropping)

我們可以使用自動裁切避免掉前兩節所述方法 (低能量映像點之移除、每列等數量移除低能量映像點) 的缺點：形狀破壞及 visual coherence；自動裁切的操作法是由原影像中選擇一個子視窗，它的尺寸是目標影像的大小，內容包含了原影像高能量的部分，也就是預期上圖中重要之物的部分。但若重要之物分散在圖中各處，此方法將難以實行。

(d) 整行性移除低能量映像點

整行性移除低能量映像點，效果介於移除映像點與裁切之間；由於移除映像點與裁切各有優缺點，我們可以選擇折衷的方法，優點方面雖不如先前所述方法來得完備、有效率，但亦可減少先前所述方法在缺點方面的限制及明顯破壞。然而，使用「整行性移除低能量映像點」，成果影像仍可能出現不自然的加工人造感。

2.2 接縫雕刻 (Seam Carving)

先前所述方法：低能量映像點之移除 (remove the pixels with lowest energy)、每列等數量移除低能量映像點 (remove an equal number of low energy pixels from every row)、自動裁切 (Auto-cropping)、整行性移除低能量映像點 (remove whole columns with the lowest energy) 皆有其不可避免之缺點，我們需要一個更好的影像尺寸重製工具，它必須不像自動裁切及整行性移除低能量映像點那麼受限制，又能比低能量映像點之移除、每列等數量移除低能量映像點來得能夠保護影像的內容，如此一來成果影像才合於普遍性的運用；根據以上的需求，我們採用接縫雕刻 (seam carving) 這個方法。

所謂接縫是由影像中低能量映像點所連接起來的連續線，可以由上到下，或是由左到右。接縫的最優性 (optimality) 是由影像能量函數 (image energy function) 所決定。重複地在一個方向執行插入或移除接縫的動作，我們可以改變影像的縱橫比；若在縱、橫兩個方向都執行插入或移除接縫的動作，我們可以重定影像的尺寸 (retarget the image to a new size)。執行影像尺寸重製時接縫的選擇和排序，是由能量函數 (energy function) 決定的，能

量函數設計的考量是要保護影像中的重要之物，而重要之物通常是邊緣 (edge) 較多的部分，所以不論是要放大或縮小影像，我們所選擇將移除或是插入的接縫，都是經過較少邊緣、能量較低的部分。前節所提之能量 (energy)，又可以看作是接縫的成本 (cost)。假設 I 是一張大小為 $n \times m$ 的影像，則垂直接縫 (vertical seam) 的數學定義為：

$$\begin{aligned} S^x &= \{s_i^x\}_{i=1}^n \\ &= \{(x(i), i)\}_{i=1}^n, \text{ s. t. } \forall i, |x(i) \\ &\quad - x(i-1)| \leq 1 \end{aligned} \quad (2)$$

式(2)中的 \mathcal{X} 是一個對映 (mapping) 函數，它的範圍為 $[1, \dots, m]$ 。該式的意思是：這條垂直接縫是一條由上到下、寬為 1 個映像點，且是在八鄰居 (8-neighbor) 範圍內的連續路徑。水平接縫 (horizontal seam) 的概念亦同，差別在水平接縫是由左到右。

所謂接縫雕刻演算法，就是將所找到的能量最小的垂直接縫 (vertical seam) 或水平接縫 (horizontal seam) 移除，如圖 1 所示，並且補償被移除的部分；若要縮小兩個映像點，則再重新操作此演算法一次，以此類推。接縫雕刻是一種影像操作技術，支援內容感知影像尺寸重製 (content-aware image resizing)，定義上雖放在移除水平或垂直接縫以縮小影像，但將接縫雕刻的概念延伸，事實上是可以放大及縮小影像；除此之外，接縫雕刻亦可應用於影像重點內容突出 (image content enhancement) 及影像物件移除 (object removal)。接縫雕刻使用能量函數 (energy function) 來定義影像中各個映像點的重要程度；前一小節提到，接縫是一條由低能量映像點所連接起來的連續路徑，從左到右或由上到下穿越影像；經由成功地移除或插入合適的接縫，我們便可以縮小或是放大影像的尺寸，並且是在兩個方向都可以完成。在接縫雕刻裡，我們希望移除較多低能量映像點及較少的高能量映像點，藉此保護預期中影像的重要內容之處，也就是在大部分的情況下，邊緣較多、變化較多的部分。

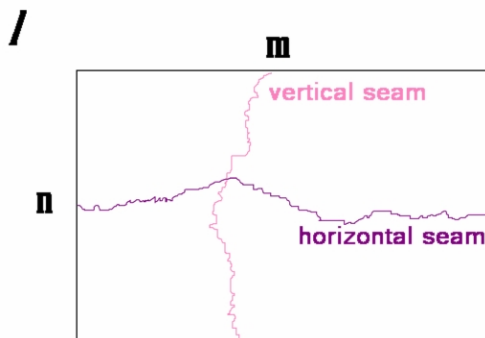


圖 1 垂直接縫及水平接縫示意圖

先前提到能量函數 (energy function) 設計的考量是要保護影像中的重要之物，而重要之物通常是邊緣 (edge) 較多的部分。因此能量函數是以影像中邊緣多寡、變化程度為基礎而設計產生；邊緣較多的部分，因大部分的狀況下為影像中重要內容部分，透過能量函數予以保護；反之，邊緣較少的部分則透過能量函數選擇為可移除或可複製的部分。

然而，並非所有影像都能循此規則進行內容感知影像尺寸重製 (content-aware image resizing)，經過反覆實驗發現，背景複雜的影像，其背景被偵測到的邊緣很可能比前景重要之物的來得多，如此一來能量函數、接縫雕刻系統的設計反而有違欲操作影像的特性，如圖 5-7；除了背景，若影像中較不重要的部分邊緣多且與重要主體邊緣數量落差大亦同，如圖 5-8。所以接縫雕刻的限制在於若影像的背景、不重要的部分過於複雜、邊緣多且和重要主體的邊緣數量落差大，將可能產生不良的結果。

為了解決此問題，本論文以一般接縫雕刻為基礎，加上人臉偵測及人身保護的功能，克服原本在一般接縫雕刻上容易失敗的例子。加上人臉偵測功能之後，影像處理過程中，會先將人臉判斷出來並對人臉及人身加以保護，如此一來可以保留住我們所認為影像中較重要的人身部分，相較之下較不重要的、視覺上較不容易注意到的地方則加以移除。實驗結果顯示，所提演算法於人像影像中能夠更加準確保留人們所想要保留的部分，有別與以往的「內容感知影像尺寸重製」僅是以邊緣多寡來判斷影像中的重要部分，使用內容感知影像尺寸重製及人臉與人身保護的合併方法所處理的影像可以更加的貼近我們的需求。

3. 所提演算法

如前面所言，傳統接縫雕刻之限制為，影像背景若較為複雜，其背景所偵測到的邊緣很有可能比前景重要之物來得多，因此會有影像重要之物 (例如：人像) 產生不可接受之變形的可能，這樣的結果對我們來說是不甚理想的。本論文以接縫雕刻內容感知影像尺寸重製為基礎，加入人臉偵測與人身保護的功能。在影像處理過程中，先對影像進行膚色偵測，偵測出膚色的區域，再利用橢圓遮罩框出人臉部分，隨後調高人像臉部的能量值 (energy value)，使接縫 (seam) 不會從人臉及人身的部分穿過，進而達成保護人臉及人像的目標，詳細介紹如下。

3.1 人臉偵測之方法

人類的皮膚顏色在影像中，若以 RGB 三原色彩空間來看，人眼可以輕易的偵測出膚色的區域，但對電腦並不是件容易的事情，且在 RGB 色彩空間中，易受到光源變動的影響，因此我們利用對光照變化較不敏感的顏色空間—HSV 進行分割，其中 HSV 分別代表，色度 (Hue)、飽和度

(Saturation)、亮度 (Value)。

然而針對不同光源環境下對膚色作特定的閾值切割，將膚色區域從影像中分離出來，必須將設定的閾值做一個調整。因此需要搭配類神經網路配合影像處理技術進行膚色學習，以較為客觀的方式將膚色區域從背景中分辨出來，如圖 2 所示。



圖 2 影像的膚色區域隔離 (a)原始影像；(b)影像經由膚色隔離的膚色區域。

3.2 橢圓遮罩搜尋法與人臉及人身保護

前一小節說明了如何找出膚色的部位，而由於人臉外型的邊界近似橢圓形，所以運用了橢圓遮罩來框住膚色的區域，保護人臉不因影像縮小或放大遭到破壞。

橢圓是由它的長短軸來決定形狀及大小，由於人臉的大小與攝影機跟目標物的距離有關，所以橢圓遮罩必須配合膚色區域改變大小，而人臉所形成橢圓的長短軸比例約 1.2 : 1，按照這比例可以縮小搜尋的空間。其數學模型如下，我們定義搜尋狀態 m 如下所示：

$$m = (x, y, \sigma) \quad (3)$$

搜尋空間為：

$$M = \{m: |x - x_0| \leq x_r, |y - y_0| \leq y_r, |\sigma - \sigma_0| \leq \sigma_r\} \quad (4)$$

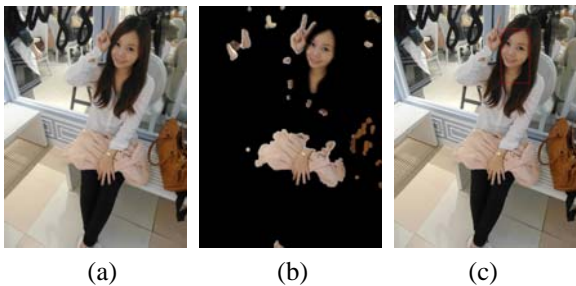


圖 3 影像的人臉偵測 (a)原始影像；(b)影像經由膚色隔離的膚色區域；(c)經由程式框出人臉部分。

其中 x_r 跟 y_r 為在點 (x_0, y_0) 時的搜尋範圍， σ_r 代表橢圓半徑搜尋範圍， σ_0 為圓心定義的短軸半徑長度。結果如圖 3 所示。

在找出橢圓遮罩之後，將橢圓內之映象點的能量值調高，使原本通過此區域的接縫計算出來的能量值變高，而不被計算為最佳接縫，如此在執行移除或加入接縫時，人臉部分區域便不會出現變形和破壞，詳細實驗結果如圖 4 所示。

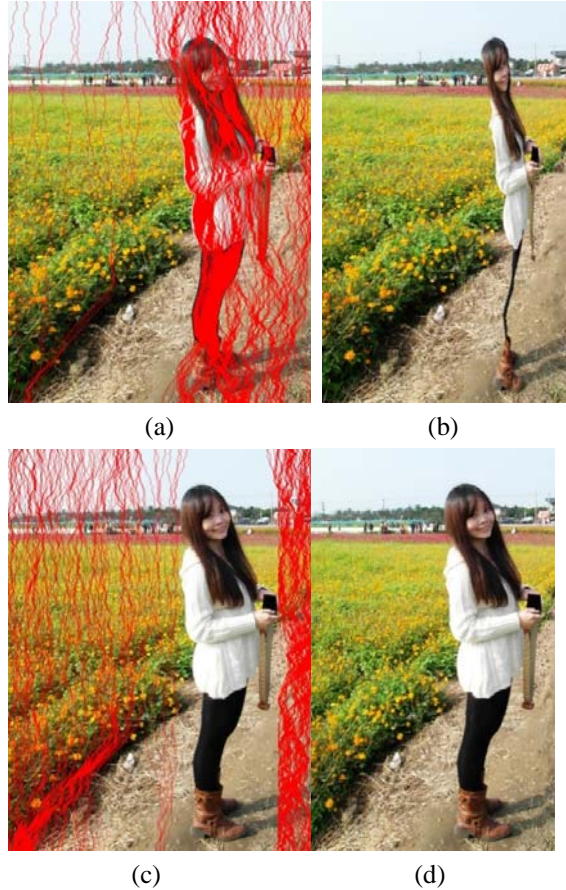


圖 4 一般接縫雕刻與加入人臉偵測人身保護之接縫雕刻兩成果比較；(a)加入人臉偵測與人身保護前，經由接縫雕刻程式所選擇的接縫；(b)加入人臉偵測與人身保護前，經由接縫雕刻縮小影像寬度之成果；(c)所提加入人臉偵測與人身保護後，經由接縫雕刻程式所選擇的接縫；(d)所提加入人臉偵測與人身保護後，經由接縫雕刻縮小影像寬度之成果。

4. 實驗結果

在實驗部分我們就一般非等比例影像縮放、自動裁切 (Auto-cropping)、一般接縫雕刻 (seam carving) 及所提加入人臉偵測與人身保護之接縫雕刻演算法的成果做評估與比較。

經數次實驗後發現，一般非等比例影像縮放在影像中較不重要部分的變形可能可以接受，但影像中重要主體亦會發生變形，這樣的變形會使得成果影像並不合於使用，尤其在主體為人物的影像中更

是如此，如圖 5(a)；自動裁切雖保護了影像中的重要主體，但原影像勢必會有物件被部分或全部性地移除，原影像包含的訊息會有遺漏，它所要傳遞的訊息在成果影像並不一定會順利傳遞，若被移除的物件為不重要的物件或原先就希望能移除的物件，這樣的結果是可接受的，但若被移除的物件包含我們希望透過影像傳遞的訊息，或者重要物件散佈影像各處，自動裁切亦不合於使用，於是我們認為自動裁切不適用於各種影像，此方法對輸入影像的限制仍高，如圖 5(b)。

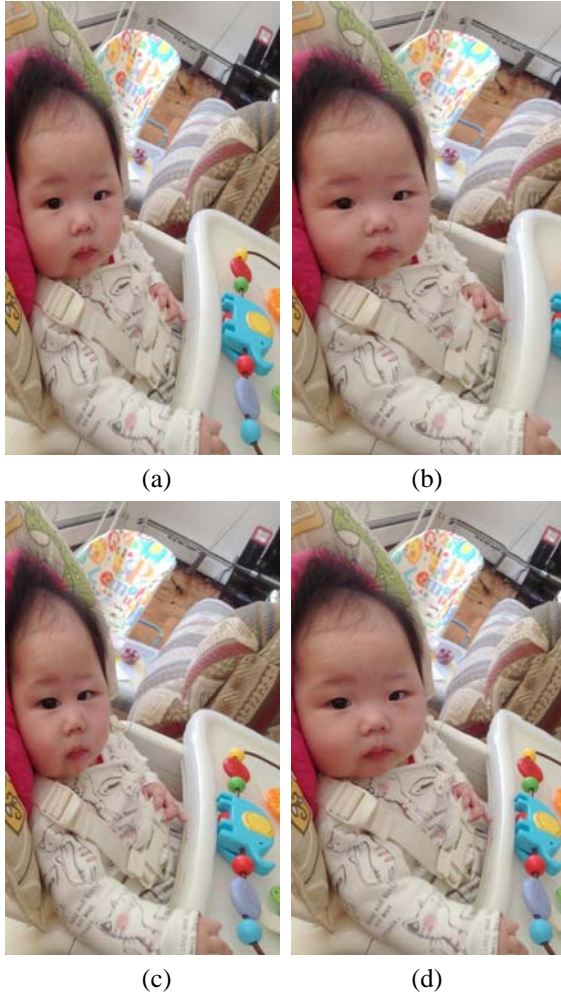


圖 5 各方法之縮小影像寬度成果比較；(a)使用一般非等比例縮放縮小寬度至 80%之影像成果；(b)使用自動裁切縮小寬度至 80%之影像成果；(c)經由傳統接縫雕刻縮小寬度至 80%之影像成果；(d)所提加入人臉偵測與人身保護後，經由接縫雕刻縮小寬度至 80%之影像成果。

若我們希望傳達的訊息是寶寶與面前的玩具，在自動裁切的成果影像中將只有寶寶，面前的玩具已被移除，我們無法完整接收欲傳達的訊息。前面提到一般接縫雕刻於背景複雜的影像將會出現重要主體變形的失敗成果，原影像是一張背景複雜的影像，成果如圖 5(c)，的確出現重要主體的人

像變型問題，該結果並無法被採納與使用；最後我們將看到本專題的研究成果，加入人臉偵測與人身保護之接縫雕刻，如圖 5(d)，影像重要的人像部份，也就是寶寶的部份，完整、無變形地保留下來，而寶寶面前的玩具亦能有效展示。在第五章的時候提到，影像在各行各業都扮演重要角色，假設這張寶寶照片是一間嬰兒玩具廠商委託設計與印刷業者為其設計的平面廣告使用影像，為了影像顯示裝置的問題，必須將原圖的寬度縮小至 80%，那麼可以看到，圖 5(d)的所提加入人臉偵測與人身保護之接縫雕刻的成果影像最合於使用，完整地展示了寶寶與玩具。

圖 6 與圖 7 所示為另外兩張人像測試影像，其同樣顯示所提演算法確實能比傳統接縫雕刻演算法能有更好的結果。

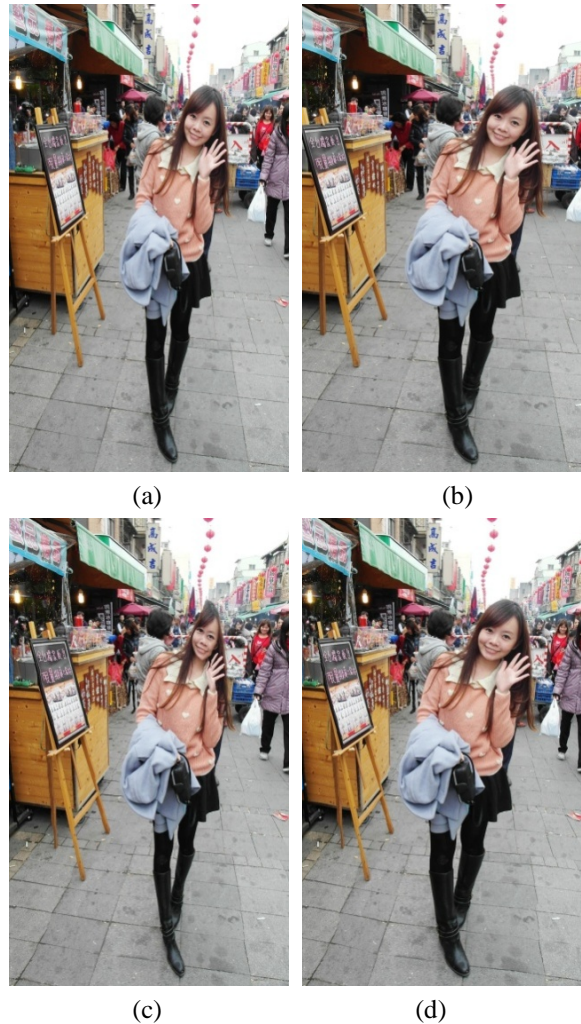


圖 6 各方法之縮小影像寬度成果比較；(a)使用一般非等比例縮放縮小寬度至 80%之影像成果；(b)使用自動裁切縮小寬度至 80%之影像成果；(c)經由傳統接縫雕刻縮小寬度至 80%之影像成果；(d)所提加入人臉偵測與人身保護的接縫雕刻縮小寬度至 80%之影像成果。



圖 7 各方法之縮小影像寬度成果比較；(a)非等比例縮放縮小影像寬度之成果；(b)自動裁切縮小影像寬度之成果；(c)加入人臉偵測與人身保護前，經由接縫雕刻縮小影像寬度之成果；(d)加入人臉偵測與人身保護後，經由接縫雕刻縮小影像寬度之成果。

5. 結論

科技一直在進步，各個和科技有關的領域也蓬勃發展，不僅是學術上的進步、使用，人們的生活也和科技環環相扣、緊密聯繫。尤其是影像方面，人們如今每天都會看到各式各樣的影像，不論是路上看到的平面廣告、電腦裡存的相片、透過網路傳送的及時新聞影像，亦或是手裡拿的智慧型手機，我們和影像已有了密不可分的關係。

一般非等比例影像縮放，雖然突破了「等比例」的限制，更有彈性，但影像重要主體變形使得此方法並不合於廣泛使用。傳統非等比例影像縮放，本專題報告所提之四方法，各有其優點但也各有其無法避免之缺點，彈性提升但限制仍在。若要突破破壞影像內容此重大限制，我們必須採用「內容感知影像尺寸重製」(content-aware image resizing)，接縫雕刻(seam carving)可以於大部分的影像上順利完成保護影像重要主體的任務，但經反覆實驗發

現，若輸入之影像背景複雜或影像中較不重要的部分邊緣(edge)多且與重要主體邊緣數量落差大，程式將會誤判影像重要主體為不重要、可變形或可刪除的部分，接縫反而都從重要主體經過，導致縮放影像後，重要主體變形的不可接受成果。

各種影像中，又以人像影像最普遍，也最不可接受重要主體變形，在本研究中，我們改善輸入人像影像的接縫雕刻程式，我們採取人臉偵測後做人身保護，成果明顯改善許多。但若重要主體佔影像的比例過大，可操作的空間便會明顯被壓縮，即使做人臉偵測及人身保護改善後的接縫雕刻，成果能夠改善的幅度依然很有限，可以把解決這個問題當作未來的研究目標。

參考文獻

- [1] Shai Avidan and Ariel Shamir. "Seam Carving for Content-Aware Image Resizing." ACM Transactions on Graphics Volume 26, Number 3, SIGGRAPH 2007.
- [2] Tsu-Hsu Liu. "Improving Seam Carving for Content-Aware Image Resizing Using Saliency Map." 2010.
- [3] Tsung-Yen Chen. "Design and implementation of the face tracking and recognition system." 2007.
- [4] El-Kjamy, S.-E., Hadhoud, M.-M., Dessouky, M.-I., Salam, B.-M., and El-samie, F.-E.-A., "A new edge preserving pixel-by-pixel (PBP) cubic image interpolation approach," Proceeding of National Radion Science Conference, March 2004, PP. C11-1-9.
- [5] Weiming Dong, Ning Zhou, Jean-Claude Paul and Xiaopeng Zhang. "Optimized image resizing using seam carving and scaling." 2009. Proc of 11th ACM Workshop on Multimedia and Security. Princeton, New Jersey.
- [6] Lukasz A. Stasiak and Raul Vicente-Garcia. "Identity recognition-based correction mechanism for face tracking." AMDO'10 Proceedings of the 6th international conference of Articulated motion and deformable objects, 2010, pp. 203-212.
- [7] Young-Ouk Kim, Joonki Paik, Jingu Heo, Andreas Koschan, Besma Abidi and Mongi Abidi. "Automatic Face Region Tracking for Highly Accurate Face Recognition in Unconstrained Environments." Proceeding of IEEE Conference on Advanced Video and Signal Based Surveillance, 2003, pp.29.