

3D Photography(2/2) 2005/5/25

B90902015 呂旺洲

B90902084 潘振銘

B90902086 黃尹聖

大綱:

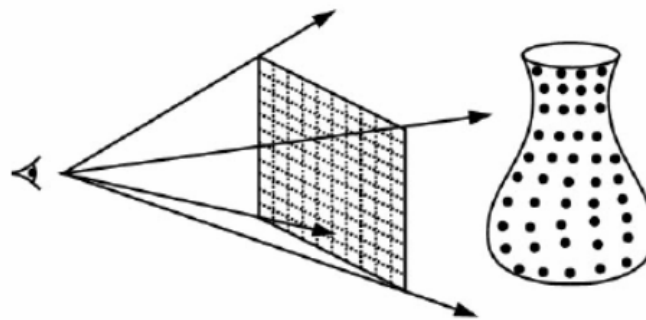
一開始會對上星期講的 Range Acquisition 做一個簡單的 Review，接著講如何利用這些 Range Image 來建立一個完整的 3D Model。基本上需要兩件事情，第一件事情是如何把不同的 View Align 到同一個 space 下面，這邊講方法叫 ICP。第二件事情是我們如何把 Align 好的 3D Image 建成一個 3D Model，今天要介紹的方法是 Volumetric reconstruction。最後介紹完這些之後，我們會來介紹一些 3D Photography 的系統 or Project。

Pipeline:

Range Acquisition Review -> Align 3D Image 的方法(ICP) -> 把 3D Image 結合成 3D model 的方法(Volumetric reconstruction) -> 3D Photography 的應用專題。

Range Acquisition

所謂的 Range Image 就是從某個 Camera View 拍出去的 Image 和一般 Image 不同的是每個 Pixel 存的是 Z value(Depth 從相機到物體的距離)，如下圖



Range image

如何得到 Range Image?

可以用 Contact，斷層掃描的方式或是利用物體反射的特性，用光，聲波，電磁波的反射得到與物體的距離(Depth)。我們會 Focus 在 Optical 的方式上面，在 Optical 方面又可以分成三類，Active，passive 和 Active-variant of passive

Passive Method:直接用照片分析，而不把訊號射到場景裡面。

Active Method:射入訊號到場景內再用其反射的特性分析。

Active-variant of passive Method:理論上看起來是 passive method 不過實際上還是投入了某些訊號。

Passive Method:

Stereo: 利用人眼的原理，兩個 view 的 disparity，求出物體的 Depth。

Space Carving: 用 Multi-view reconstruction 的方式

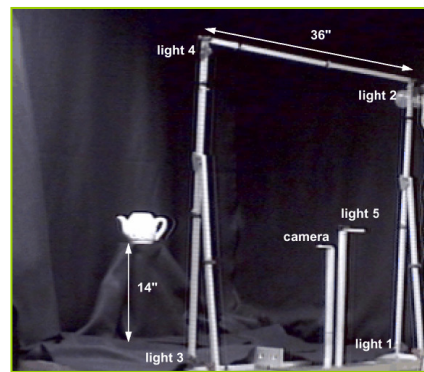
Active Method:

基本上利用 Triangulation，讓我們知道說投影的光的平面在哪裡，從兩個像機射各出一條線各經過兩 Image Plane 上同一物體的 Pixel 這兩條線的交點就是物體在 3D 中的位置。如左下方這張圖是 Cyberware 的 Whole body scanner. 如果沒有那麼多預算也可以利用簡單的 Shadow scanning 只需要一個 Web-cam，一盞燈和一根筷子。(右下圖)



Active variants Method:

投入不同位置的光源,去看物體在不同光源下的成像是怎樣子,它不用假設物體是 diffused(傳統 stereo 方法都要求是 diffused),它沒辦法直接求出物體的 Depth,它是看物體的 Normal 來反推出物體的 Depth.如下圖



3D Model Pipeline

- 1) View Planning: 在拍 Range Image 時,要規劃使的每個角落都拍到
 - 2) 3D Scanner: 用上面講的方法取得 Range Image
 - 3) Alignment: 用接下來要講的 ICP 方法,把 Range Image Align
 - 4) Merging: 把 Align 好的 3D Image 建成 3D Model.
- (ps.今天要講的重點是 Alignment & Merging)

【ICP】(Iterative Closet Points)

ICP的基本假設是，correspondence point 就是該點距離最近的那個點(Iterative Closet Points 的 Closet Points 就是這樣來的)。這個假設很有可能是錯的，可是無所謂，我們可以根據這個假設找到 optimal transformation 去做 align，基本上，這個 align 不會做得很好，因為假設可能是錯的，不過我們可以根據新做出來的 mesh 再做同樣的假設，一直 iterative 地做下去，一直到整個系統不再動了為止。在大部份的情況下，通常會收斂。

接下來介紹 ICP 可能的變形。

根據 ICP 演算法可以分成六個步驟，每一個步驟分別可以做改進，這些步驟分別是：

- **Selecting** source points (from one or both meshes)
- **Matching** to points in the other mesh
- **Weighting** the correspondences
- **Rejecting** certain (outlier) point pairs
- **Assigning** an error metric to the current transform
- **Minimizing** the error metric

<1> **Selecting** source points (from one or both meshes)

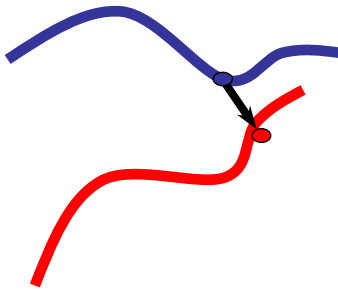
就是如何在兩個 mesh 上找出一些點當作 correspondence 的點。

- (1) All points：把所有點都丟進去用，好處就是會比較準確，壞處是會很慢
- (2) Uniform sampling：因為用的是 range image，例如就每 3X3 的 page 做 sample。
- (3) Random sampling：譬如我要取 1000 個點，就 random 的在 mesh 上選出 1000 個點。
- (4) Normal-space sampling：range image 每一點除了對應的位置外，還有它的 normal，normal 基本上就是 range image 的 gradient。我們要每個方向的 normal 都有差不多個數的點，這樣可以避免一種情形，像是如果一個 mesh 是一個平面，上面只有一個凹下去的十字，用以上三種方法，通常不會選到十字裡面的點，這樣變成對 2 個平面做 alignment，結果一般都不好，不會收斂到正確的 solution。所以這種方法也是 paper 推薦的方式。

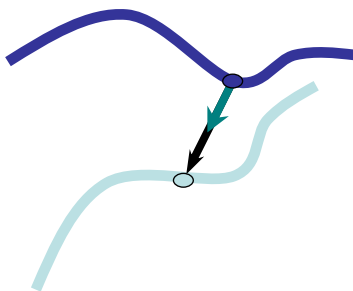
<2> Matching to points in the other mesh

原來的 ICP 是選 Closet Point 當作 correspondence point，但是有人提出了其他的選點方法。這個是影響收斂速度最大的地方。以下 4 種方法，以(4) Projection 花的時間最少，因為其他三個都要用 k-d tree 來找，要花比較多的時間。

(1) Closet Point：找距離最近的點。由於要建 k-d tree，所以速度不快。

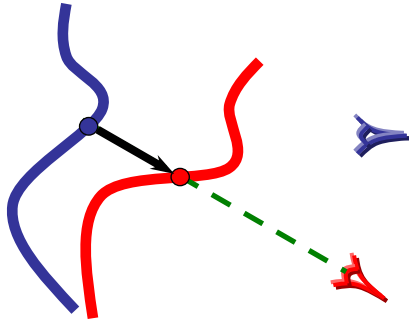


(2) Normal Shooting：除了最近以外，兩個點的 normal 不能差太多，不然就八成是選錯點了。方法是由藍色的點(如下圖)的 normal 延長交另一 mesh 於一點，即是我們想要的點。對於 smooth 的 mesh 會比第一個方法好，但對於複雜和 noise 較多的 mesh 就不太適用。



(3) Closet Compatible Point：就是以上兩種方法的加強，除了 normal 不能差太多外，color 也不能差太多，然後可以利用 feature matching。

(4) Projection：因為 input 是兩張 range image，基本上我們可以得到 camera 的座標系。如下圖，我們可以把藍色和紅色放到同一個座標系裡然後可以找到藍色的點會投影到紅色的哪一個點。這樣的方法，只需要乘一個矩陣而不需要做 search，雖然在每一個 iteration 中不一定是收斂最快的，但是所花時間是最少的，所以整體來說，Projection 這個方法是收斂速度最快的。paper 也建議使用這種方法。

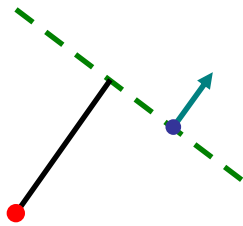


<3> **Weighting** the correspondences & **Rejecting** certain (outlier) point pairs

之前提到的 error function: $E = \sum_i^N \|p_i - (Rq_i + t)\|^2$ 是假設每個點 weight 都是 1, 那就有人想辦法要怎麼對每個點做不一樣的 weight 或 reject 某些太誇張的點。在 range image 中, 基本上 normal 愈朝向你的點會愈準, 愈偏移的會愈不準, 所以可以利用這個對每一個點做不一樣的 weighting 讓結果會更好。rejection 就是之前說的 outlier, 在經過 transformation 後, 如果發現兩個點距離相差太遠, 那就幹脆不要理它。不過 paper 裡說這個部份對於 performance 影響不大。

<4> **Assigning** an error metric to the current transform

我們怎麼定義兩點最近的距離? 一個是直接算兩個點在 3D 空間中的距離, 另一個是我們知道每的點的 normal, 而 normal 和點會形成一個平面, 所以兩點間的距離可以當作是一點到另一點所形成平面的距離。point-to-plane 的方法會比 point-to-point 收斂的速度要快。



<5> **Minimizing** the error metric

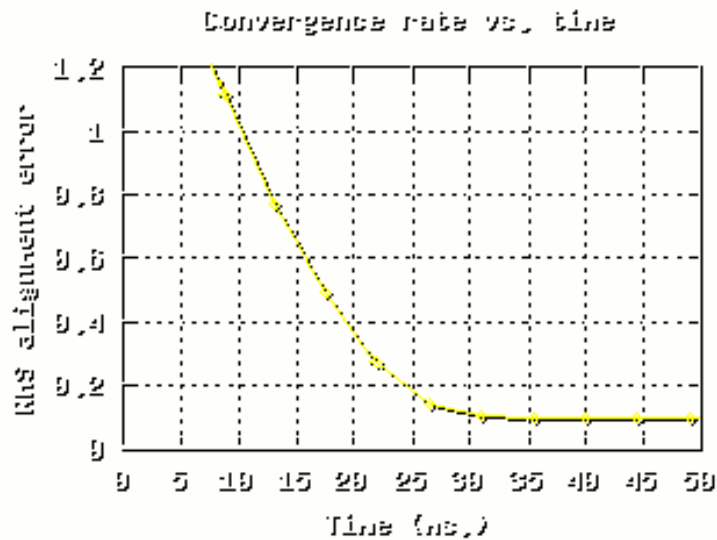
基本上沒什麼變化, 就找一個最好的 minimize algorithm 就行了。

有一篇 paper 有比較這幾種 component 對 performance 的影響, 基本上, 怎麼做其實都不會差太多, 只有少數幾個 component 會有比較大的影響, 這篇 paper 的建議是:

- (1) Selecting 的時候用 Normal-space sampling
- (2) Matching point 的時候用 Projection 的方式
- (3) Weighting 和 Rejection 不用在乎

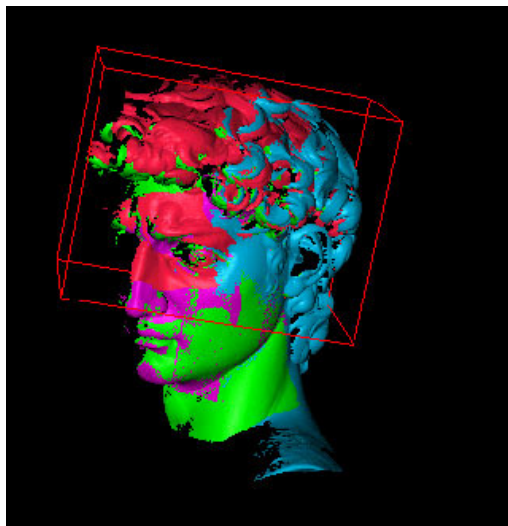
(4) Assigning error metric 用 point-to-plane distance 會比較好

然後這篇 paper 最後利用上面的方法提出了一個 high-speed ICP algorithm，做 alignment 只需要 30 個 ms，一般的話需要約 1 秒左右



所以，基本上整個流程就是：

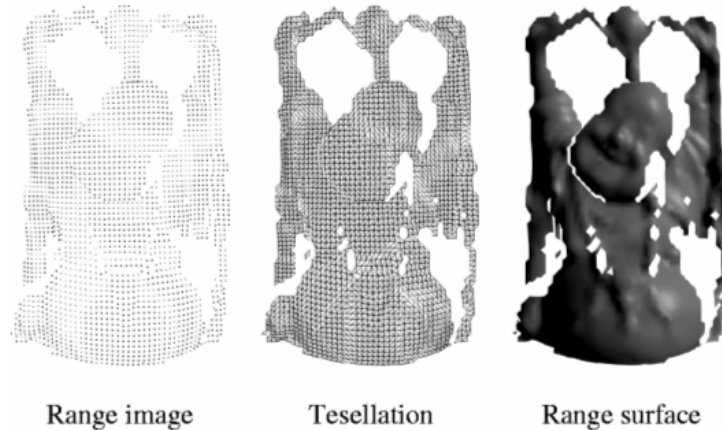
1. manual initial alignment
2. ICP to one existing scan
3. automatic ICP of all overlapping pairs
4. global relaxation to spread out error
5. merging using volumetric method



Volumetric Reconstruction(Goal:把 Align 好的 surface 建成 3D model)

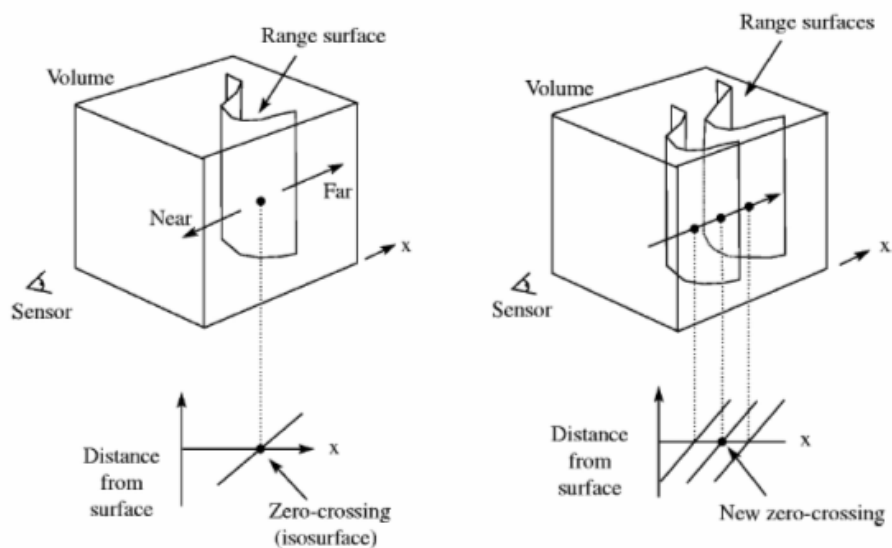
1)產生 Input Mesh

在用 ICP 把 Range Image Align 好之後接著就要把 Range Image 合起來變成一個 3D Model。要用到的方法叫做 Volumetric reconstruction。這個方法要求 Input 要是一個 surface or Mesh(網)，所以要把 Range Image 在 3D 中相鄰的點相連形成 triangular mesh，形成 Input 就是一大堆 Align 好的 Range Surface。如下圖



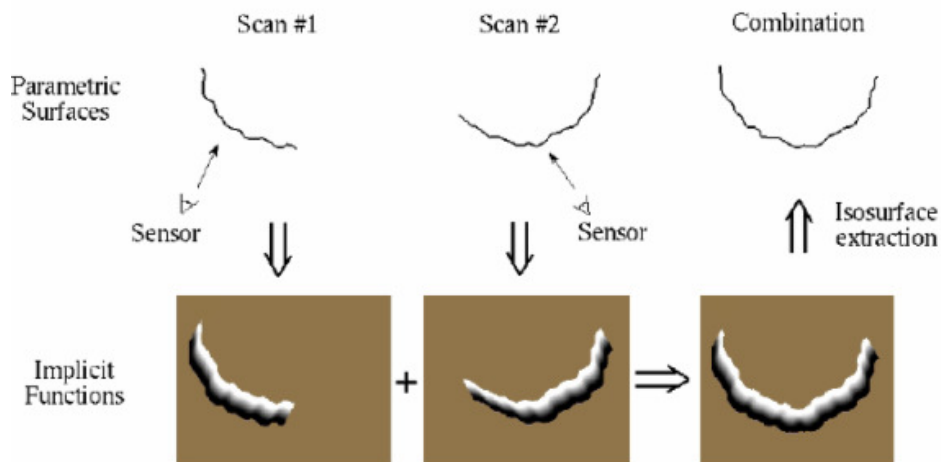
2)建 Volume & 轉到同座標系 & Signed Distance Function

在 3 度空間中建一個立方體(volume)，使所有的 Range Surface 都是在 Volume 的某一個位置，這時座標系就已經固定在這個 Space(volume)，而取得各個 Range Surface 的 Sensor 位置也轉移到這個座標系。在這些前提下我們可以對每個 Range Surface 定義 Signed Distance Function 其對任一點 X 的值就是沿著 Sensor 到 X 的射線上，Range Surface 到 X 的距離。它除了值還有正負號，以 Surface 為界，靠近 Sensor 為負(object 外面)，遠離 Sensor 為正(object 裡面)。如下圖



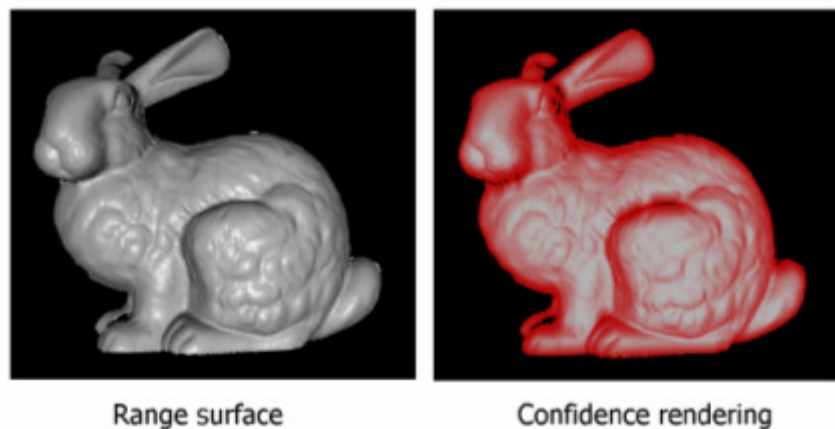
3)Combine Signed Distance Function

假如我們現在有兩個 Surface 要結合，那就有兩個 Signed Distance Function，我們把 Signed Distance Function 為零的部份叫做 Iso-surface。要結合最簡單的方法就是把兩個 Signed Distance Function 平均，形成新的 Signed Distance Function 它為零的地方就是結合的 surface 叫做 Iso-surface。如下圖



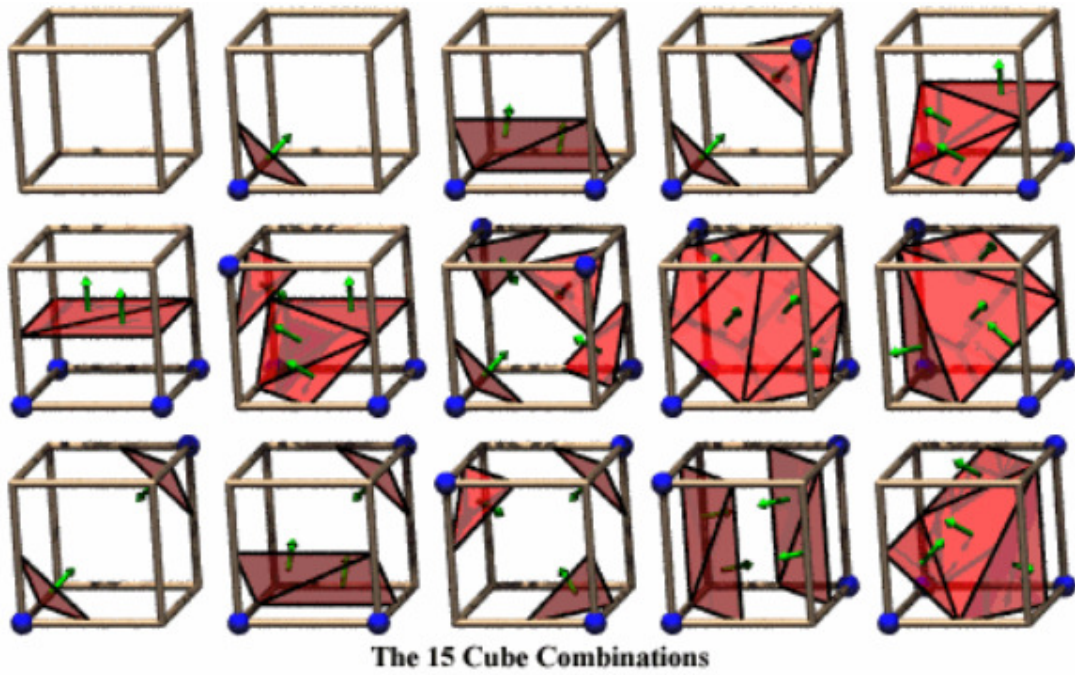
4)Weighting

我們在做 Signed Distance Function 相加平均時，還可以幫他加上 Weighting，因為當物體表面的 Normal 朝向我們時或是靠 Range Image 中心，測量會比較準確。這時就給他比較高的 Weight。如下圖兔子身上越白 weight 越高，愈紅 Weight 越低。



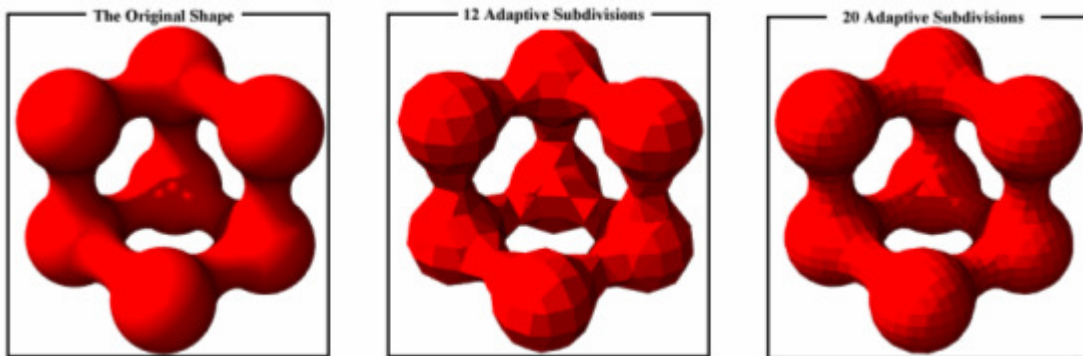
5)取出 Iso-surface

在 Combine 全部的 Signed Distance Function 後，接著就要把它的 Iso-surface 取出來，我們用到一個演算法叫 Matching Cube，在每個 Cube 的角上面有正有負，在 Signed Distance Function 的前提下，我們可以推測 Iso-surface 在哪通過，依著 Cube 頂點的正負號可以分成下面圖中 15 個 Case(藍色的頂點表示為異號)，就可找出 Iso-surface.



6) Cube 解析度

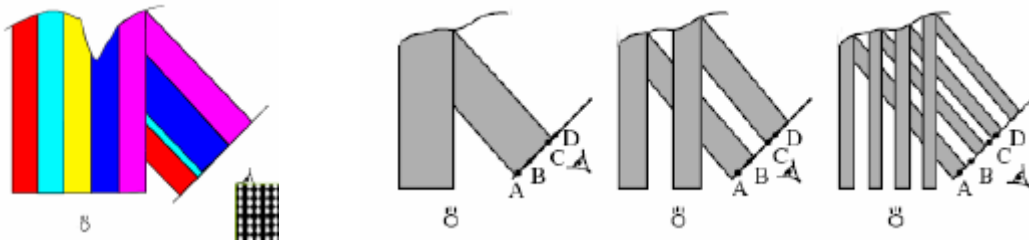
隨著 Cube 切的越細，結果也越精細，如下面圖左到右是原圖，低解析，高解析的 Match Cube。



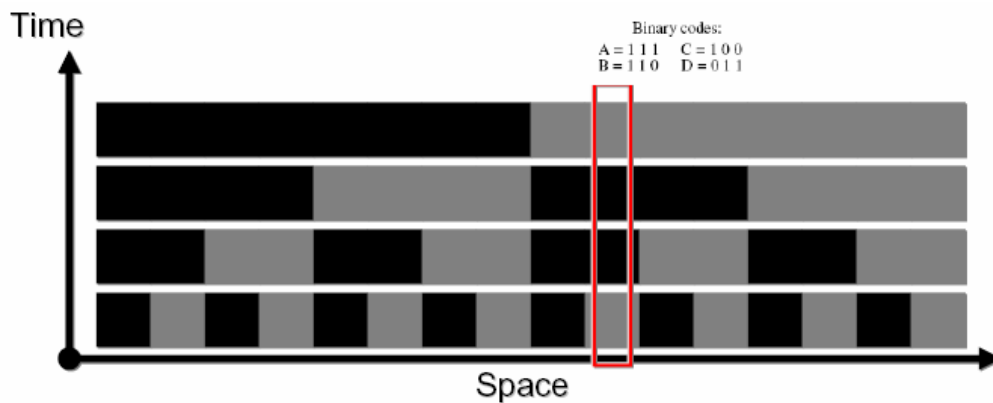
Systems, projects and applications

◆ Real-time 3D model acquisition 包括以下幾個部份：

1. View Planning：原本爲了取得完整model而不會有破洞，要決定好要從哪些角度拍攝物體，不過由於是real-time因此可以一邊拍攝一邊調整。
2. 3D Scanner：
 - Faster acquisition： 投射複數條射線至物體，方法可能有time-coded或color-coded。
 - single-stripe method原本需要N張圖片，但multiple-stripes time-coded method僅需logN張,大幅降低拍攝時間。



※color-coded and time-coded light patterns

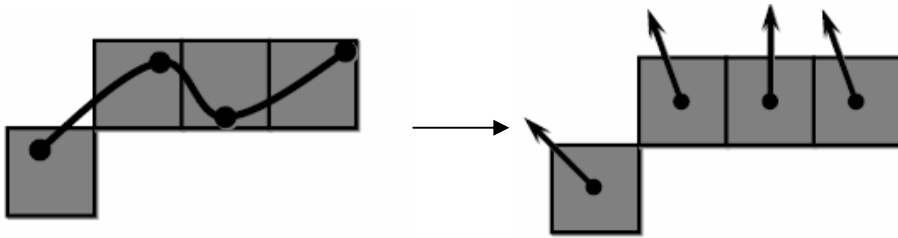


※time-coded acquisition

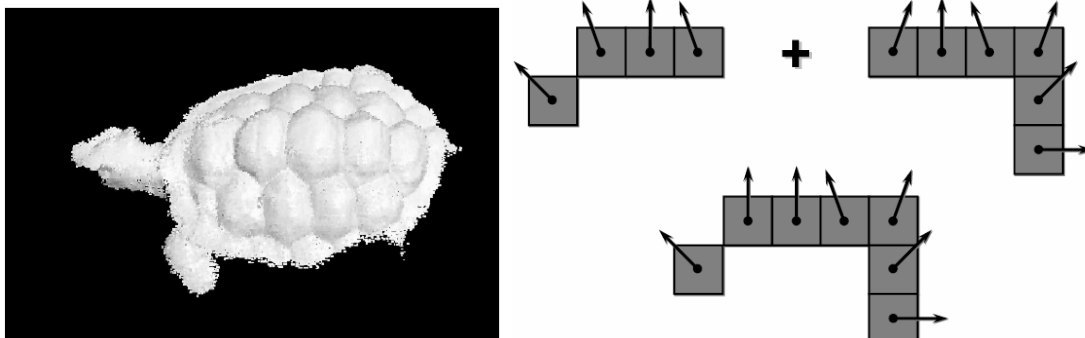
- 使用frame-to-frame tracking來決定相關的邊界。



3. Alignment : Fast ICP Algorithm ◦
4. Merging : 不用原來的曲線表示而改以normal表示，以求快速計算。



5. Display : 使用point rendering, accumulated normals for lighting ◦



※point rendering

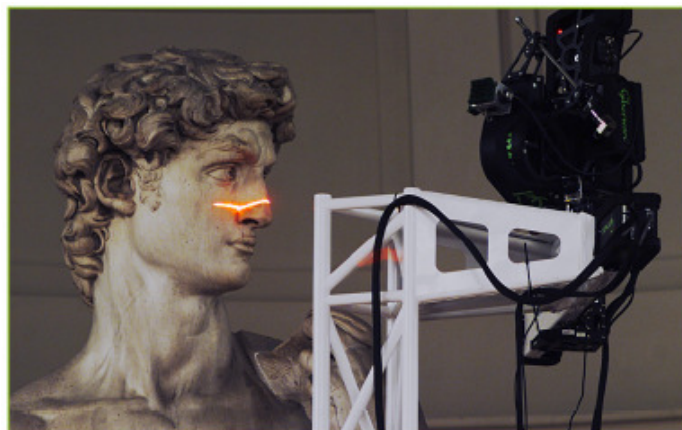
※accumulated normals for lighting

◆ The Digital Michelangelo Project :

由Stanford University的Marc Levoy主導，目標是掃描10個由米開朗基羅所雕刻的雕像，並且能夠取得高解析度的model，未來可供後人研究當初米開朗基羅是用怎樣的雕刻刀、怎樣的彫刻法才能雕出這些雕像。

掃描過程大致可分為下面幾項：

1. manual initial alignment
2. ICP to one existing scan
3. automatic ICP of all overlapping pairs
4. global relaxation to spread out error
5. merging using volumetric method

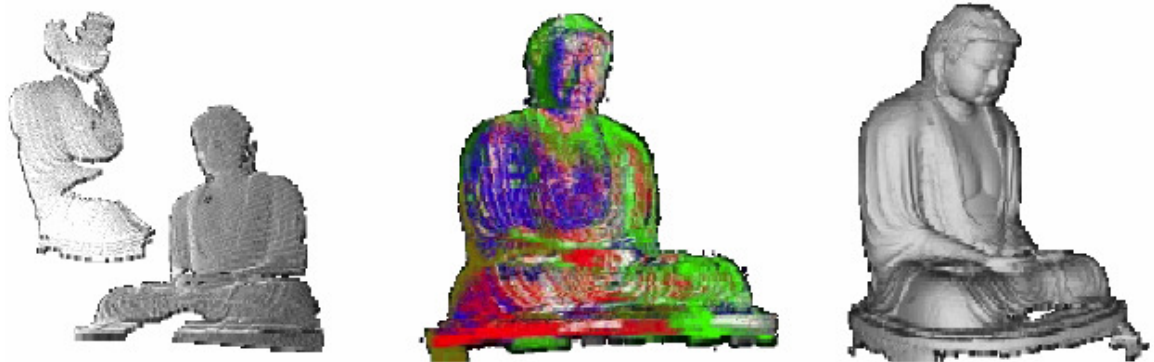


◆ **The Great Buddha Project :**

由東京大學Katsushi Ikeuchi 所主導，目標是掃描鎌倉時代所留下的大佛像，以保存日本的文化遺跡。

掃描過程大致可分為下面幾項：

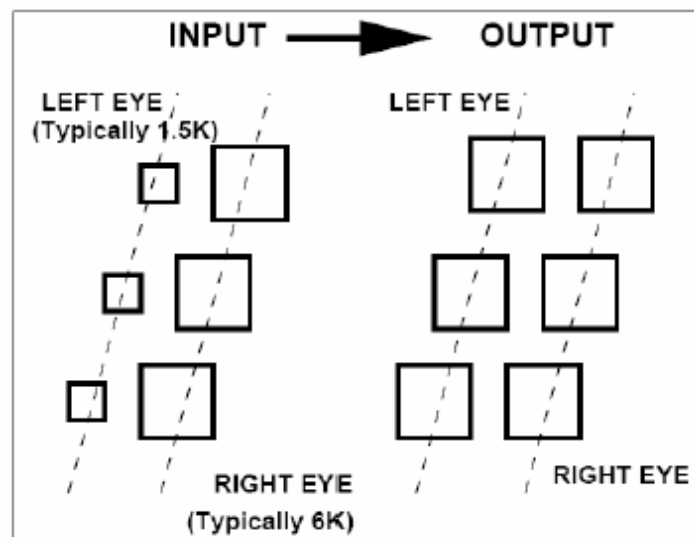
1. 20 range images (a few million points)
2. Simultaneous all-to-all ICP
3. Variant of volumetric merging (parallelized)



◆ **Hybrid stereo camera :**

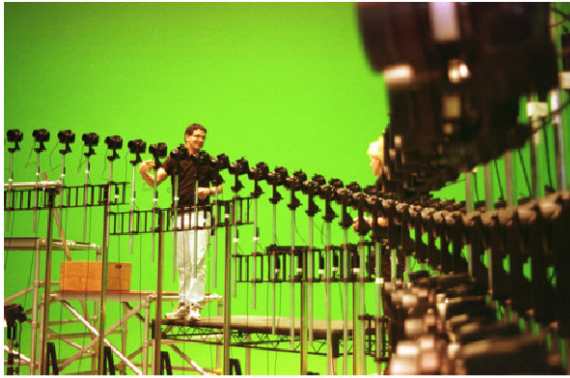
以往的IMAX 3D拍攝太過耗時，且相機爲了取得兩眼高解析度資料，往往非常的笨重，不利於live-action的拍攝。

Hybrid stereo camera用低解析度的相機(left eye 1.5k & right eye 6k)取得的資料，來獲得高解析度的資料(both left & right eye 6k)，改善相機的體積以及rendering的時間。



◆ **View interpolation :**

以往的Bullet time video是利用大量的相機擺設在固定的路徑上，由相機的照片組成每frame的影片，但如此做來不但耗費大量金錢，而且由於路徑固定，很沒有彈性，因此有了利用少數攝影機做view interpolation的研究，讓使用者可以一邊觀賞影片一邊決定自己喜歡看的角度。



Bullet time video



High-quality video view interpolation