

## 第二章 基礎理論介紹

本章的主要旨在於介紹 MPEG 的基礎理論與 SIP 協定的基礎原理。MPEG 的編碼理論中，會簡單介紹 Interframe Coding 與 Intraframe Coding、移動估測與移動補償、區塊比對等基礎理論。SIP 協定的基礎原理會介紹 SIP 的簡介、元件架構、主要功能、以及 SIP 基本呼叫的建立，以便我們對本篇論文有初步的認識。

### 2.1 MPEG 的編碼理論

MPEG 的編碼理論，我們會針對 MPEG 簡介、Intraframe 與 Interframe Coding、移動估測與移動補償、區塊比對及 MPEG Video Coding 五個重點進行討論。

#### 2.1.1 MPEG 簡介

MPEG(Moving Picture Expert Group)原指一個研究制定視訊、影音編碼標準的組織，後來泛指這個組織所制定出的視訊編碼標準。這個組織從 1988 年組成，到目前已經制定了 MPEG-1、MPEG-2、MPEG-4 等多項標準。[1]

- (1) MPEG-1 : MPEG 組織把 MPEG-1 定義為“在儲存媒體上儲存和擷取影像與音訊的標準” MPEG-1 為 MPEG 組織所訂定的第一個視訊壓縮標準。Video CD 與 MP3(MPGE-1 Layer 3)為 MPEG-1 最廣泛的應用。

(2) MPEG-2 : MPEG-2 為數位電視的標準，主要被應用於 DVB、DVD 等技術中。MPEG-2 提供 720\*480(Pixel)與 1280\*480(Pixel)兩種解析度。

(3) MPEG-4 : MPEG-4 擴展 MPEG-1 與 MPEG-2 標準，使 MPEG-4 支援聲音與視訊物件(Audio/Video Object)編碼、低位元率編碼(Low Bit Rate Encoding)。MPEG-4 在視訊會議、IP 網路攝影機的應用相當廣泛。

MPEG 的影像序列(Video Sequence)中，由 GOP(Group Of Picture)、Picture、Slice、Macroblock 以及 Block 組成。影像序列的階層關係就如下圖 2.1 所示，一個 GOP 可細分為多個 Picture；一個 Picture 可以細分為多個 Slice；一個 Slice 可細分為多個 Macroblock；一個 Macroblock 又可細分為 4 個 8\*8 pixel 的 Block。

(1) GOP : GOP 由許多 Picture 組成，通常由 I-Frame 開始，接著跟隨著 P-Frame 與 B-Frame。這三種 Frame 我們將稍後做詳細的介紹。

(2) Picture : Picture 即為影像序列(Video Sequence)中的 Frame。

(3) Slice : Slice 為 Picture 中的一部分，一個 Slice 又由整數個 Macroblock 所組成。

(4) Macroblock : Macroblock 是 16\*16(Pixel)的區塊，Macroblock 是作為區塊比對(Block Matching)的基本單位。

(5) Block : Block 為  $8 \times 8$ (Pixel)的區塊，Block 是 DCT 轉換的基本單位。

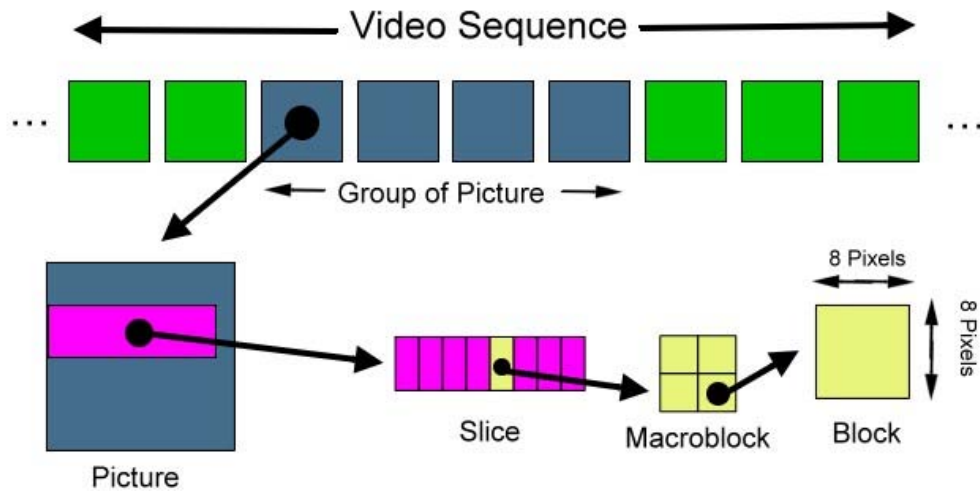


圖 2.1

MPEG 的影像壓縮，主要利用動態影片具有以下兩點特質而達到壓縮的目的[5]。

- (1) 空間資料冗餘(Spatial Redundancy)：利用人類視覺系統的特性，降低視覺系統無法辨識的影像資訊，以達到壓縮的效果。Intraframe Coding 可以減除空間資料冗餘的問題。
- (2) 時間資料冗餘(Time Redundancy)：一般的影片每秒鐘約有 20 到 30 個 Frame，因此連續的畫面中，相同位置的影像資訊必定有相當的關聯性，利用紀錄連續影像的差異也可以達到壓縮的效果。Interframe Coding 就是要利用連續影像的相關性，減除時間冗餘的問題。

## 2.1.2 Intraframe 與 Interframe 編碼

MPEG 影像壓縮的過程中，編碼的方式主要有兩種[5]，Intraframe Coding 與 Interframe Coding。Intraframe Coding 不會參考其他的影像資訊，所以 Intraframe 也可說是靜止影像(Still Image)的編碼。Interframe Coding 則必須參考其他的影像資訊進行編碼，Interframe Coding 需要進行移動估測(Motion Estimation)與移動補償(Motion Compensation)以達到利用連續畫面的相關性，進而得到壓縮的效果。

Intraframe Coding 主要是減少單張影像中空間冗餘(Spatial Redundancy)的資料。壓縮的過程以  $8*8$ (Pixel)的區塊(Block)為基本單位，進行編碼轉換中的餘弦轉換(DCT, Discrete Cosine Transform)處理。經過餘弦轉換後的影像由空間訊號轉為頻率訊號，頻率訊號可以區分為高、中與低頻訊號，其中影像的資訊大部分集中於低頻訊號，而且人類對中、高頻訊號不如低頻訊號敏感，因此可以藉由餘弦轉換後，去除不必要的高頻訊號，以達到壓縮的效果。因為 Intraframe Coding 是透過空間訊號轉頻率訊號進行壓縮，所以並不需要任何參考影像。

Interframe Coding 則利用連續的影像畫面，必定有高度關聯的特性，達到壓縮的效果。Interframe Coding 會將時間相鄰的影像畫面進行比較，並且只紀錄相鄰影像的差值，而不需要紀錄全部的影像資訊。換句話說，Interframe Coding 需要參考影像。在 MPEG 的編碼法則中，Interframe Coding 是透過移動估測(Motion

Estimation)與移動補償(Motion Compensation)來得到相鄰影像的差值。移動估測與移動補償可說是 Interframe Coding 最重要的部分，稍後我們也將針對移動估測與移動補償進行簡單地討論。

根據 Intraframe Coding 或 Interframe Coding，MPEG 的影像序列(Video Sequence)又細分為 I-Frame(Intra Frames)、P-Frame(Prediction Frames)與 B-Frame(Bi-direction Frames)。I-Frame 屬於 Intraframe Coding，P-Frame 與 B-Frame 則為 Interframe Coding。下圖 2.2 說明 P-Frame 是利用前一張影像資訊(Previous Frame)得到目前的 P-Frame 影像。下圖 2.3 說明 B-Frame 利用前一張影像(Previous Frame)與後來的影像(Future Frame)得到目前的 B-Frame 影像。

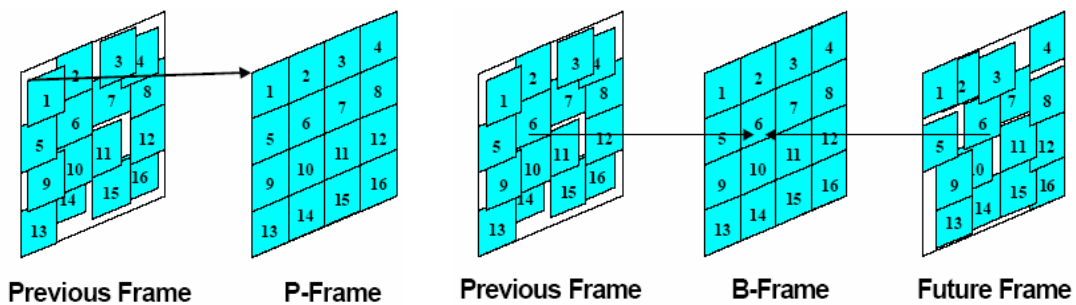


圖 2.2

圖 2.3

MPEG 的標準中，Frame 的編碼順序(Encoding Order)與播放順序(Playback Order)是不一樣的。以下圖 2.4 為例，播放的順序是由  $I_0 B_1 B_2 P_3 B_4 B_5 P_6 B_7 B_8 I_9$ ，而編碼時，B-Frame 必須參考前一張影像與後一張影像，所以需要先編碼  $P_3$ ，才

能參考  $I_0$  與  $P_3$  產生  $B_1$  與  $B_2$ ，故編碼的順序為  $I_0 P_3 B_1 B_2 P_6 B_4 B_5 I_9 B_7 B_8$ 。

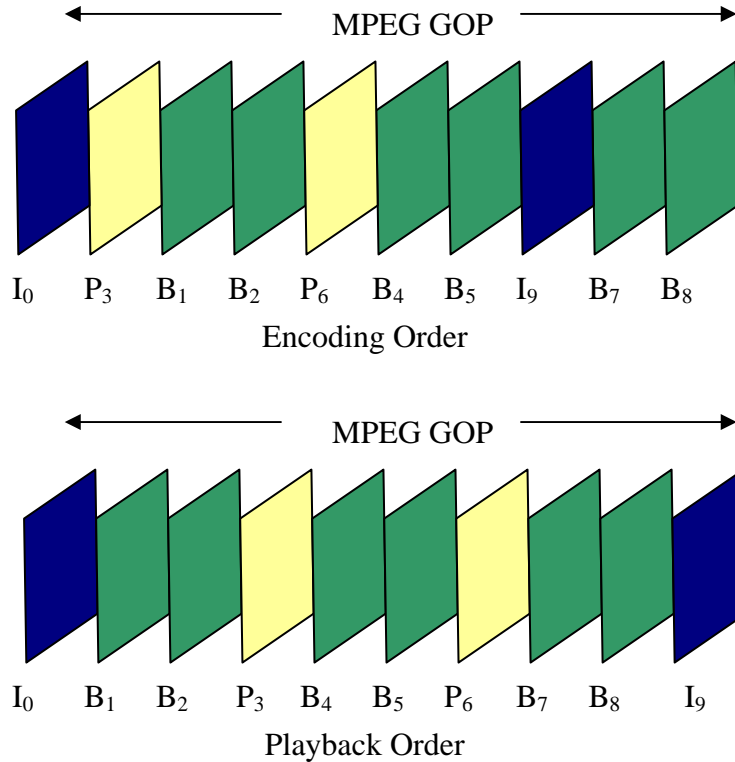


圖 2.4

### 2.1.3 移動估測與移動補償

移動估測(Motion Estimation)是利用區塊比對(Block Matching)的方法，在相鄰影像中找尋最相似的區塊。這兩個相似區塊的相對差值(Relative Distance)稱為移動向量(Motion Vector)。得到移動向量後，將要壓縮的影像區塊(即 Current Frame)與參考影像區塊(即 Previous Frame)做 Difference 運算，可以得到殘餘影像(Residual)，這個過程就稱為移動補償(Motion Compensation)。當要編碼的區塊與

參考區塊相似的程度越大，則殘餘影像的資料越少，需要紀錄的資料就越少，壓縮的效果也越好。

### 2.1.4 區塊比對

區塊比對(Block Matching)的技術中，會將目前的影像(Current Frame)分成相同大小的區塊，也稱之為來源區塊(Source Block)。每個來源區塊在參考影像(Reference Frame)中有一個對應的搜尋區域，區塊比對的目的就是要在對應的搜尋區域找出一個與來源區塊最相似的區塊(我們稱此最相似的區塊為 Candidate Block)。下圖 2.5 與 2.6 即為區塊比對的概念圖。

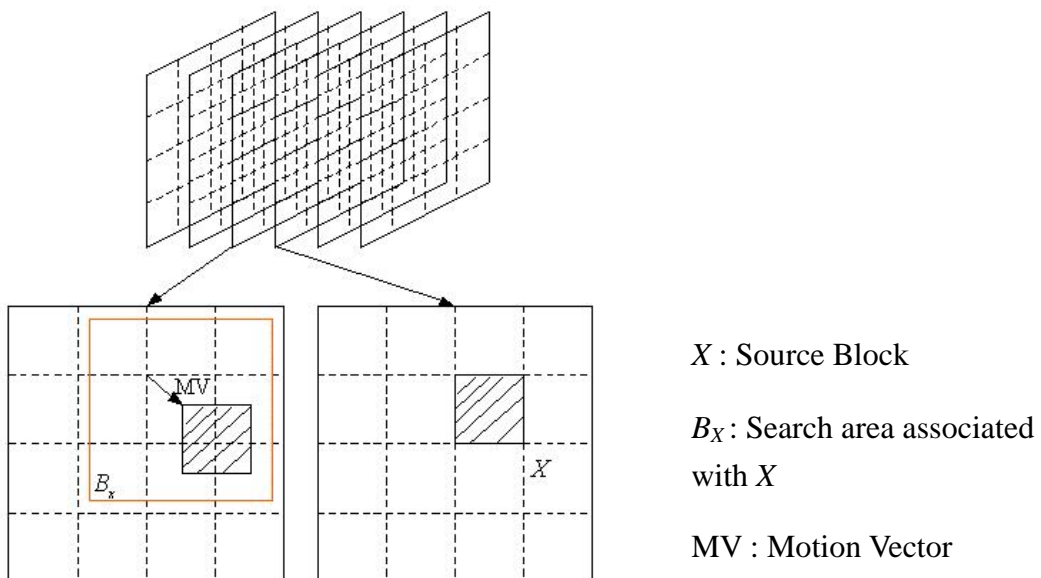


圖 2.5

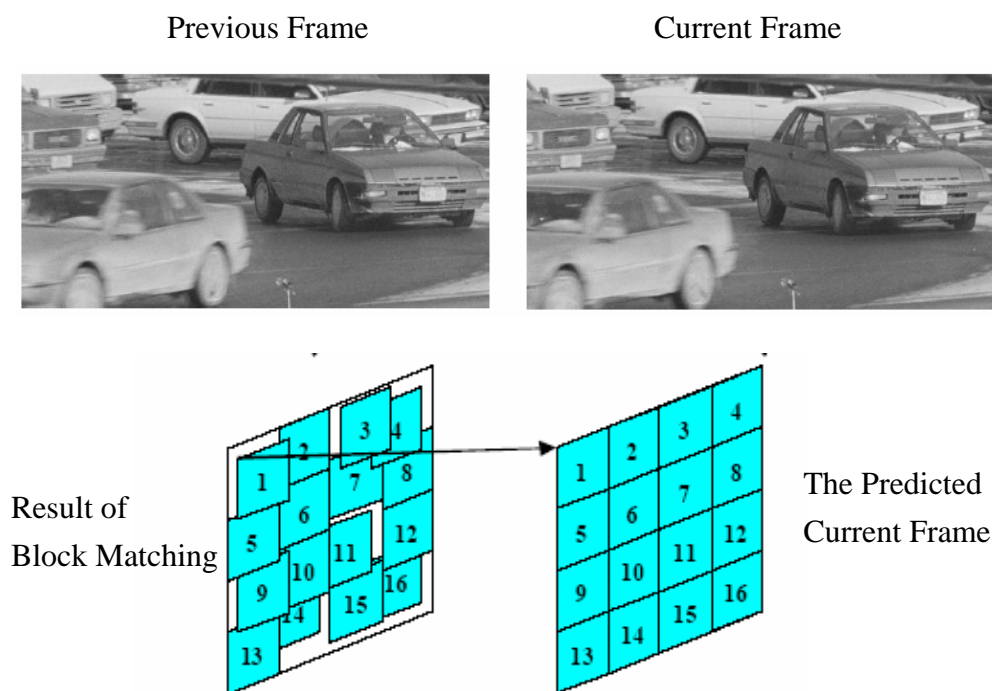


圖 2.6

區塊比對(Block Matching)的演算法主要有兩種，一種是完全搜尋法(Full Search Algorithm)，另一種為快速搜尋法(Fast Search Algorithm)。完全搜尋法可以找到最相似的區塊，換句話說，需要紀錄的資料較少，壓縮的效果比較好，但搜尋時所需花費的時間較多。快速搜尋法雖然可以降低搜尋時所花費的時間，但相對的搜尋的結果也比較不精確。

### 2.1.5 MPEG 的影像編碼

編碼器取得影像後，如果這個影像是 Intra Frame，則編碼器會將這個 Frame 進行餘弦轉換(DCT)與量化(Quantization)，最後再進行 Entropy Encoding。編碼器



進行完餘弦轉換與量化後，必須將此影像利用反量化(De-Quantization)與反餘弦轉換(IDCT)建構下一個影像的參考畫面(Reference Frame)。若此取得的影像為 Inter Frame，編碼器會先對這張影像做移動估測(Motion Estimation)，以取得移動向量(Motion Vector)。並利用所取得的移動向量進行移動補償(Motion Compensation)。進行移動補償後，會將得到的差值影像做 DCT 轉換與量化，最後再將差值影像進行 Entropy Encoding。另外，編碼器也必須將差值影像與參考影像重建，以作為下一張影像的參考影像。MPEG 的編碼的簡單概念就如下圖 2.7 所示。

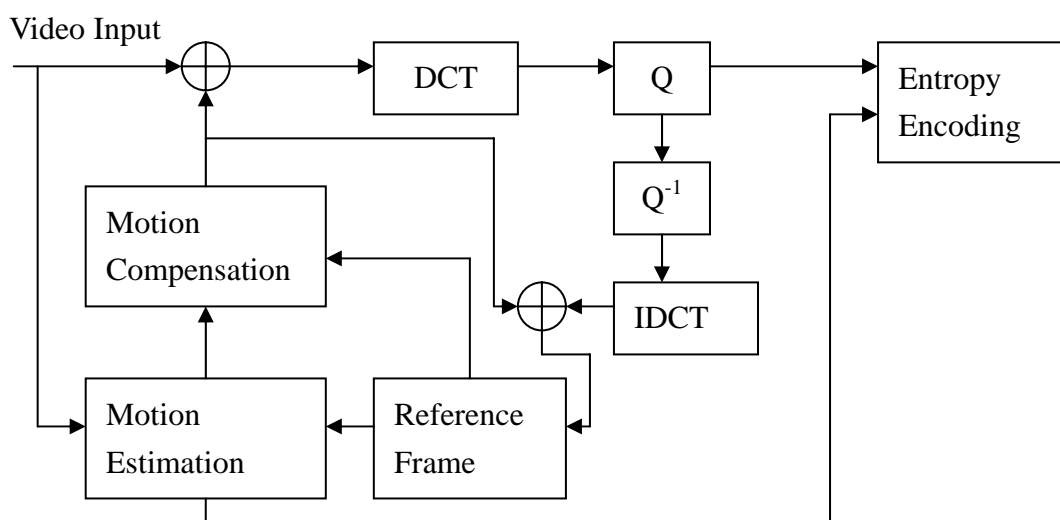


圖 2.7

## 2.2 SIP 協定基礎原理

SIP 協定基礎原理，我們會針對 SIP 的簡介、SIP 的元件架構、SIP 的主要功能、SIP 的訊息與 SIP 基本呼叫的建立做簡單的討論。

### 2.2.1 SIP 簡介

SIP(Session Initiation Protocol)是一種建立會談(session)的信號通訊協定(signaling protocol)，會談的建立可以是簡單的兩方通話，也可以是複雜的多人視訊會談。SIP 目前被廣泛的應用於網路視訊會議(video conference)、網路電話(voip phone)。

SIP 最早於 1996 年由 IETF(Internet Engineering Task Force)組織所制定，目前已將完整的規範定義於 RFC-3261 中。

### 2.2.2 SIP 的元件架構

SIP 是一種 Client 與 Server 的架構，在 SIP 的標準中定義了幾個重要的元件，這些元件中主要又分為 SIP 伺服器程式(SIP Servers)與 SIP 使用者代理程式(SIP User Agents)。

SIP 伺服器程式(SIP Servers)為整個會談(Session)控制系統，可以提供

Location、Registrar、Redirection 以及 Proxy 等功能。SIP 使用者代理程式(SIP User Agents)則被定義為會談的終端設備，負責發起 SIP Request 以建立多媒體會談(Media Session)的連線[7]。

以下為上述各元件的介紹：

1. SIP User Agents :

SIP User Agents (UA)被定義為發起 SIP Request 或者接受 SIP Request 的終端設備。主要負責多媒體會談(Media Session)的建立，以及傳送或者接收多媒體資料。UA 可以是實體的 IP 電話也可以是電腦中執行的 SIP 軟體程式。

SIP User Agents 又可以區分為 User Agent Clients (UAC)和 User Agent Servers (UAS)。

- a. User Agent Clients (UAC) : UAC 負責產生 SIP Request 以及處理 SIP Response。
- b. User Agent Servers (UAS) : UAS 負責接受 SIP Request 並且發出 SIP Response。

UA 在一個 SIP 會談(Session)中，經常需要接替扮演 UAS 與 UAC 的角色。這與一般常見的 Client and Server 的架構比較不同，例如 HTTP 的通訊

協定中，PC 端一直扮演著 Client 的角色，而 Web Server 則一直扮演著 Server 的角色。

## 2. SIP Servers:

在 RFC 所定義的 SIP Servers 分為四類，Location Server、Registrar Server、Proxy Server 以及 Redirection Server。以下，我將逐一介紹。

### a. Location Server

Location Server 當作一個資料庫使用，資料庫中紀錄使用者的資訊，如 IP 位址等。因為 SIP UA 無法直接存取 Location Server，所以 Location Server 必須將這些資訊提供給 Redirection Server 或 Proxy Server，藉以使得 UA 可以獲得被呼叫端的 IP 位址。

### b. Registrar Server

Registrar Server 負責接受 Registration Request，並且更新用戶在 Location Server 的資訊，以提供 Proxy Server 與 Redirection Server 查詢用戶的位址資訊。

c. Proxy Server

Proxy Server 主要功能在於接受 User Agent 或者其他 Server 傳送的 Request，並且將 Request 轉送到其他 Server。必要時，Proxy Server 會對 Request 的訊息重新定義。

d. Redirection Server

Redirection Server 接收到 Request 後，會依照 Location Server 所提供的資訊，將被呼叫端的新位址資訊回傳給呼叫端。使得被呼叫端即使更換了 IP 位址，也能正確的建立會談(Session)。

## 2.2.3 SIP 的主要功能

SIP 主要的功能為用戶端向系統註冊、邀請其他用戶參與會談、協商會議媒體的訊息格式、建立多媒體串流、結束會談等功能。

### 1. 用戶端向系統註冊

在 SIP 的系統中，用戶端是可能是網路中兩個點對點的裝置，有可能是電腦中的 SIP 程式，也可能是 SIP Phone 或者 PDA 等移動性很高的裝置。雖然 SIP 的會談可以透過已知的 IP 位址來建立，但隨著系統的複雜性增加，會談的建立也就跟著變的複雜，此外，當使用者在不同地方連上網際網路，它的 IP

位址資訊也會跟著改變，所以 SIP 提供了註冊的機制，讓用戶端在可以不知道被邀請的用戶端 IP 位址，也能建立會談。

SIP 提供了註冊的機制，當用戶端連上 SIP Server 時，可以傳送 REGISTER 的訊息給 Registrar Server，進行註冊的動作。當 Registrar Server 收到用戶端的 REGISTER 訊息後，會將用戶端的 ID 與 IP 位址資訊結合，並且紀錄在 Location Server 中，以提供日後其他用戶端邀請時使用。

透過註冊的機制，用戶端可以具有可移動性的優點。即使用戶端使用 PDA 在不同地方上網，具有不一樣的 IP 位址，但是只要透過註冊的功能，就可以更新用戶端在 Location Server 的資訊，即可讓用戶端可以使用相同 ID 不同 IP 位址與其他用戶端進行會談功能。

## 2. 邀請其他用戶參與會談

在 SIP 的系統中，會談的就是透過用戶端發送一個 INVITE 訊息給其他被邀請的用戶端，以建立連線，等待連線建立後，即可透過此連線傳送語音或者多媒體的訊息。

當用戶端發送一個 INVITE 訊息後，可以透過 Proxy Server 或者 Redirection

Server 向 Location Server 取得被邀請端用戶的 IP 位址資訊，並透過此資訊建立會談的連線。

### 3. 協商會議媒體格式

SIP 協定只注重傳遞的訊息內容，對於訊息內容所附加的資料並無特別的定義。所以，通常我們可以使用 Session Description Protocol (SDP,RFC 2327) 附加於 SIP 的訊息中，使得被邀請的用戶端可以了解會談的狀態資訊。

### 4. 建立多媒體串流

當被邀請的用戶端接受 INVITE 後，會談便已成功建立。在 SIP 的會談中，可以加入多媒體串流。在 SIP 中，多媒體串流與 SIP 協定訊息是分開建立的。換句話說，我們透過 SIP 協定的訊息建立 SIP 會談後，可以利用所取得的 IP 位址資訊，建立一個點對點的連線，傳送多媒體串流檔案，而不再需要透過建立 SIP 會談的 SIP Servers。

### 5. 結束會談

用戶端在 SIP 的會談結束後，可以透過 SIP 的 BYE 訊息，告知其他用戶端，進而結束這次的會談。

## 2.2.4 SIP 的訊息

SIP 的訊息類似 HTTP 的通訊協定，利用了文字的方式(Text-based)來編碼。

此外，SIP 的訊息也與 HTTP 的通訊協定一樣，可區分為 Request 和 Response 兩類。

### 2. SIP Request

SIP Request 訊息是指 Client 端傳送給 Server 端的訊息。SIP Request 又被定義出幾種不同的型態，每一種型態可以被視為 Client 端對 SIP Server 或者 SIP User Agent 所發出的不同要求或服務。

- a. INVITE：建立會談(Session)
- b. ACK：作為被呼叫端收到 INVITE 訊息的確認
- c. BYE：結束一個已存在的會談(Session)
- d. CANCEL：取消一個尚未建立的會談(Session)
- e. REGISTER：向 SIP Server 註冊用戶端的位址資訊
- f. OPTIONS：查詢 SIP Server 及其功能

### 3. SIP Response

SIP Response 訊息為 Server 端傳送給 Client 端的訊息，用來回應 Client 端所發出的 Request。SIP Response 又分成以下六大類。



a. 1xx : Informational Response

1xx Response 代表 Server 或者 Proxy 接收到 Request，並且正在處理 Request 所要求的事件。例如: 180 Ringing。

b. 2xx : Successful Response

2xx Response 代表 Client 端發出的 Request 已經成功地被處理。  
例如: 200 OK。

c. 3xx : Redirection Response

3xx Response 代表必須重新導向的回覆。例如: 302 Moved Temporarily。

d. 4xx : Request Failure Response

4xx Response 代表 SIP Server 對 User Agent Client 所發出的 Request 回覆失敗的訊息。例如: 403 Forbidden。

e. 5xx : Server Failure Response

5xx Response 代表 Server 端本身發生錯誤，而無法完成 User Agent Client 所發出的 Request。例如: 503 Server Unavailable。

## f. 6xx : Global Failure Response

6xx Response 代表 User Agent Client 所發出的 Request 在所有的 Server 皆無法被處理。例如: 600 Busy Everywhere。

## 2.2.5 SIP 基本呼叫的建立

一個簡單的 SIP 會談(Session)就如同下圖 2.8 所表示，當呼叫端(User Agent Client)想要與被呼叫端(User Agent Server)建立一個會談時，首先 User Agent Client 必須向 User Agent Server 發出一個 INVITE 的要求。User Agent Client 發出 INVITE 訊息時，這個 INVITE 的訊息會先傳送到 SIP Server，當 SIP Server 接收到 INVITE 的訊息後，會回傳一個 100 Trying 的訊息給 User Agent Client，同時搜尋 User Agent Server 所對應的 IP 位址，接著再將這個 INVITE 的訊息傳送給 User Agent Server。

User Agent Server 收到來自 SIP Server 的 INVITE 訊息後，會先回傳一個 100 Trying 的訊息，接著再回傳一個 180 Ringing 的訊息給 SIP Server 代表 User Agent Server 已收到 INVITE 的訊息，並且正在處理中。當 SIP Server 收到來自 User Agent Server 回傳的 180 Ringing 訊息，SIP Server 也會把 180 Ringing 的訊息回傳給 User Agent Client，代表 User Agent Server 已收到 User Agent Client 所發出的 INVITE 要求。

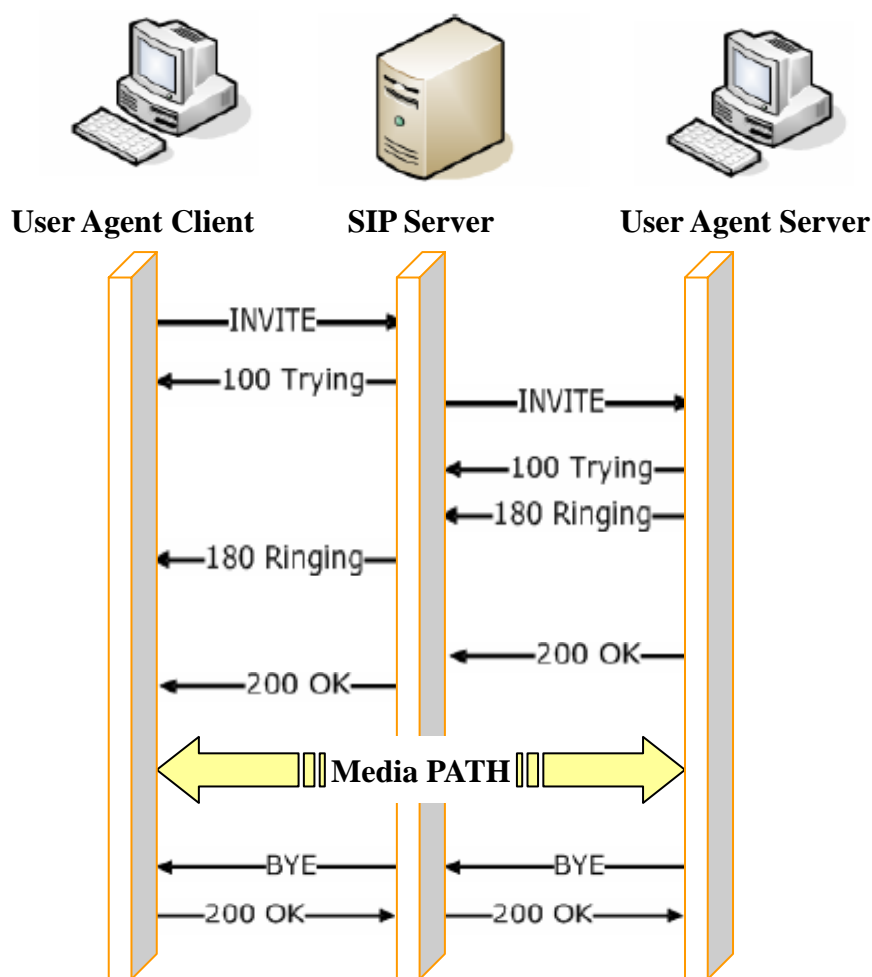


圖 2.8

User Agent Server 在確定要接受 User Agent Client 所發起的會談時，User Agent Server 會回傳一個 200 OK 的訊息給 SIP Server，代表 User Agent Server 接受這個會談的建立，SIP Server 收到 200 OK 的訊息後，也會把這個 200 OK 的訊息傳送給 User Agent Client，當 User Agent Client 收到 200 OK 的訊息時，代表這個會談已經成功建立。

會談成功建立後，User Agent Client 與 User Agent Server 就可以建立點對點的多媒體傳輸連線，這個傳輸連線的建立不再需要透過 SIP Server。User Agent Client 與 User Agent Server 可以利用這個連線傳送多媒體資料，例如影像或者聲音...等。

最後，會談的結束可有 User Agent Server 或者 User Agent Client 的任何一端發起。假設 User Agent Server 想結束這一段會談時，可以發出一個 BYE 的訊息給 User Agent Client。User Agent Server 發出的 BYE 訊息，會先傳送到 SIP Server，再藉由 SIP Server 傳送給 User Agent Client。如果 User Agent Client 接受 BYE 的要求時，就會傳送一個 200 OK 的訊息給 User Agent Server，代表 User Agent Client 同意結束這個會談。200 OK 的訊息傳送過程跟 BYE 訊息類似，也是先由 User Agent Client 傳送給 SIP Server，再藉由 SIP Server 傳送給 User Agent Server，然後結束這個會談。