



浙江大学计算机学院
数字媒体与网络技术

Digital Asset Management

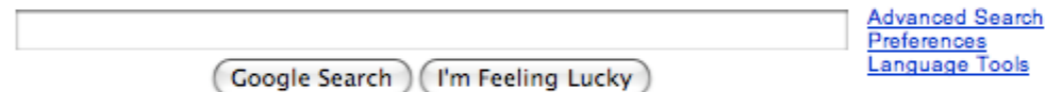
数字媒体资源管理

6. Introduction to digital media retrieval

任课老师：张宏鑫
2008-10-08

Main methods of digital media retrieval

- **Text-based** digital media retrieval

The Google logo is displayed in its characteristic multi-colored font (blue, red, yellow, blue, green, red) with a trademark symbol.A screenshot of the Google search interface from 2008. It features a search input box, a "Google Search" button, an "I'm Feeling Lucky" button, and links for "Advanced Search", "Preferences", and "Language Tools".

[Advertising Programs](#) - [Business Solutions](#) - [About Google](#) - [Go to Google China](#)

©2008 - [Privacy](#)

- **Content-based** digital media retrieval



Content based digital media retrieval

- Query by example on multimedia-data

- Demo:

- The **GNU Image-Finding Tool**

- <http://www.gnu.org/software/gif/>

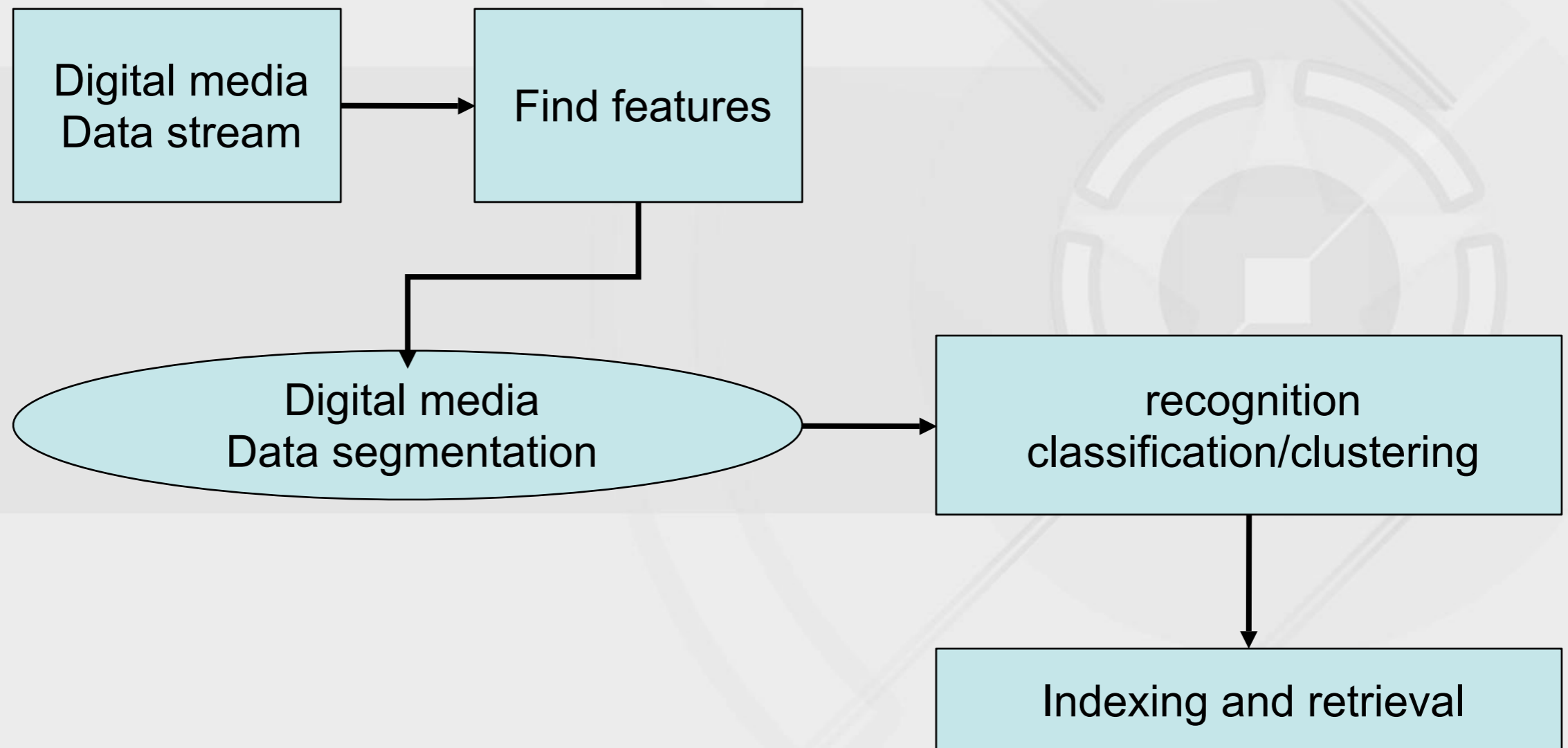


- **Flexible Image Retrieval Engine**

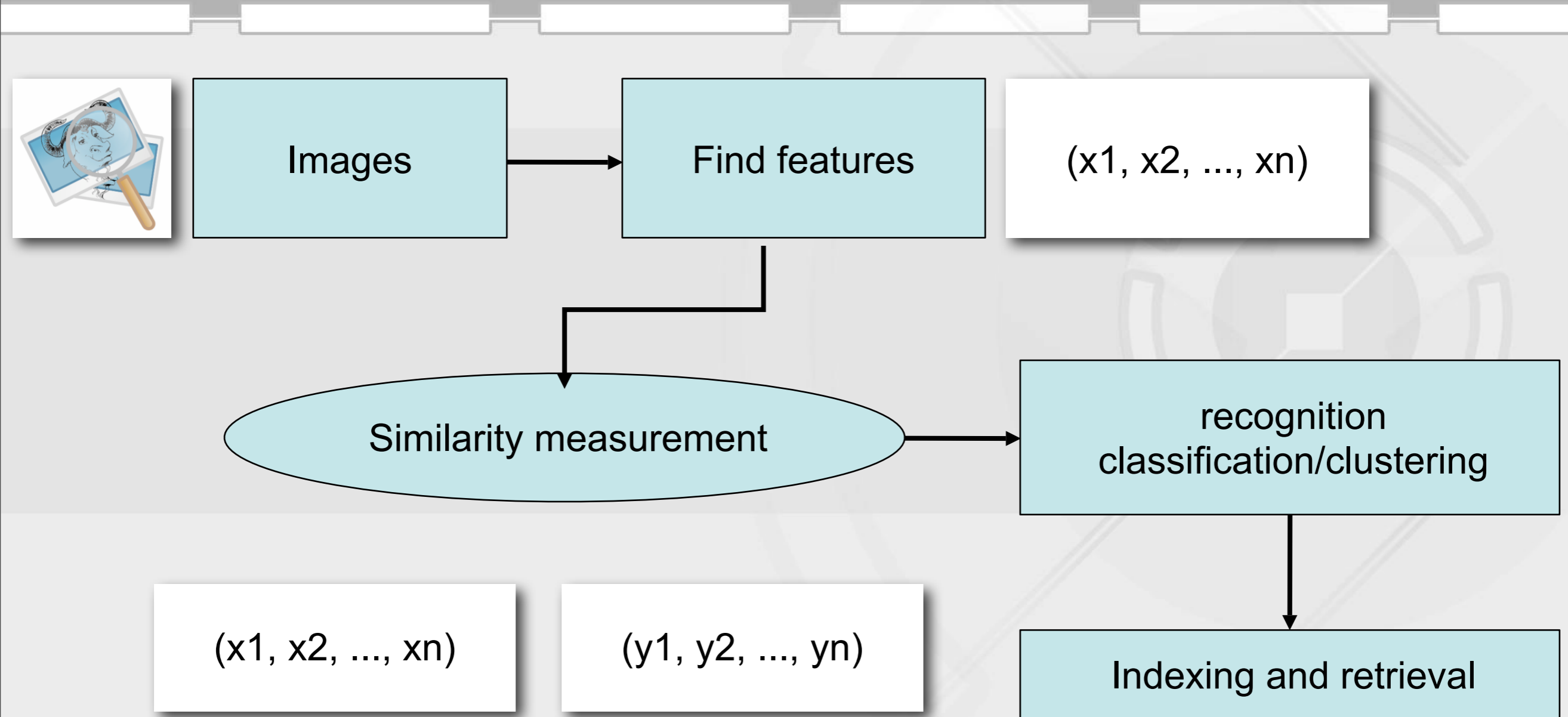
Fire



The workflow of digital media analysis and retrieval



Workflow of CBIR



Features of image

- Finding out features of image is a key step of image retrieval
 - Image-based retrieval usually need to pre-construct feature database of images for retrieval
- Major image features:
 - Color features
 - Texture features
 - Shape features
 - Space relation features



Color features of image

- Color feature is a most widely used vision feature. It is mainly used to analyze color distributions in an image, including:
 - Color histogram
 - Color moments
 - Color set
 - Color clustering vectors
 - Color relation graph

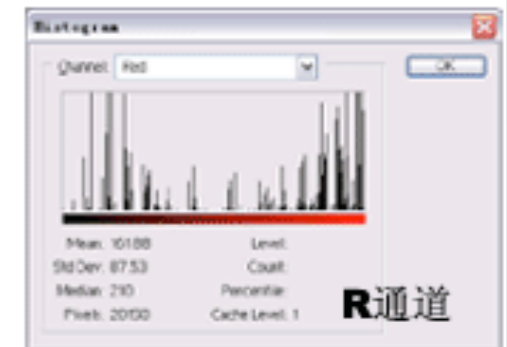
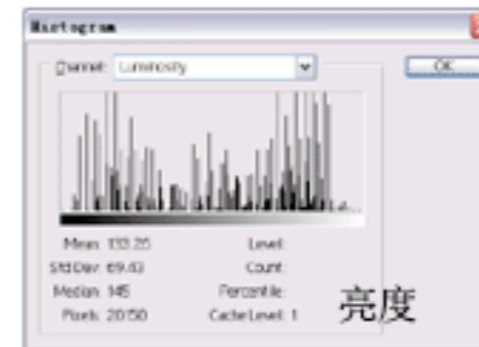


Image texture features

- Texture features are such vision features employed to measure homogeneous phenomenon in images. They are
 - independent to color or illuminance,
 - and are intrinsic features of object surfaces.
- Major texture features
 - Tamura texture features
 - Self-regression texture model
 - Transform based texture features
 - DWT, DFT, Garbor filter bank
 - others



Image shape features

- Shape features are computed out based on object segments or regions, mainly including
 - contour features
 - and regions features.
- Typical approaches include
 - Fourier shape description
 - Moment invariants



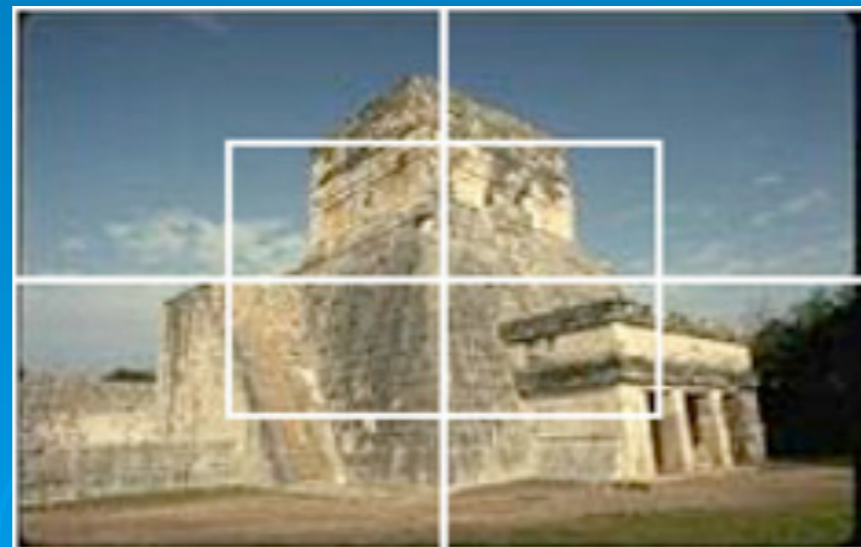
Image Retrieval Phase (cont.)

➤ Query by color anglogram (cont.)

- Convert RGB to HSV [wikipedia]



- Global and sub-image histogram forms LSI matrix.



[Zhao & Grosky 2002]



Image A

$$\longrightarrow (x_1, x_2, \dots, x_n)$$

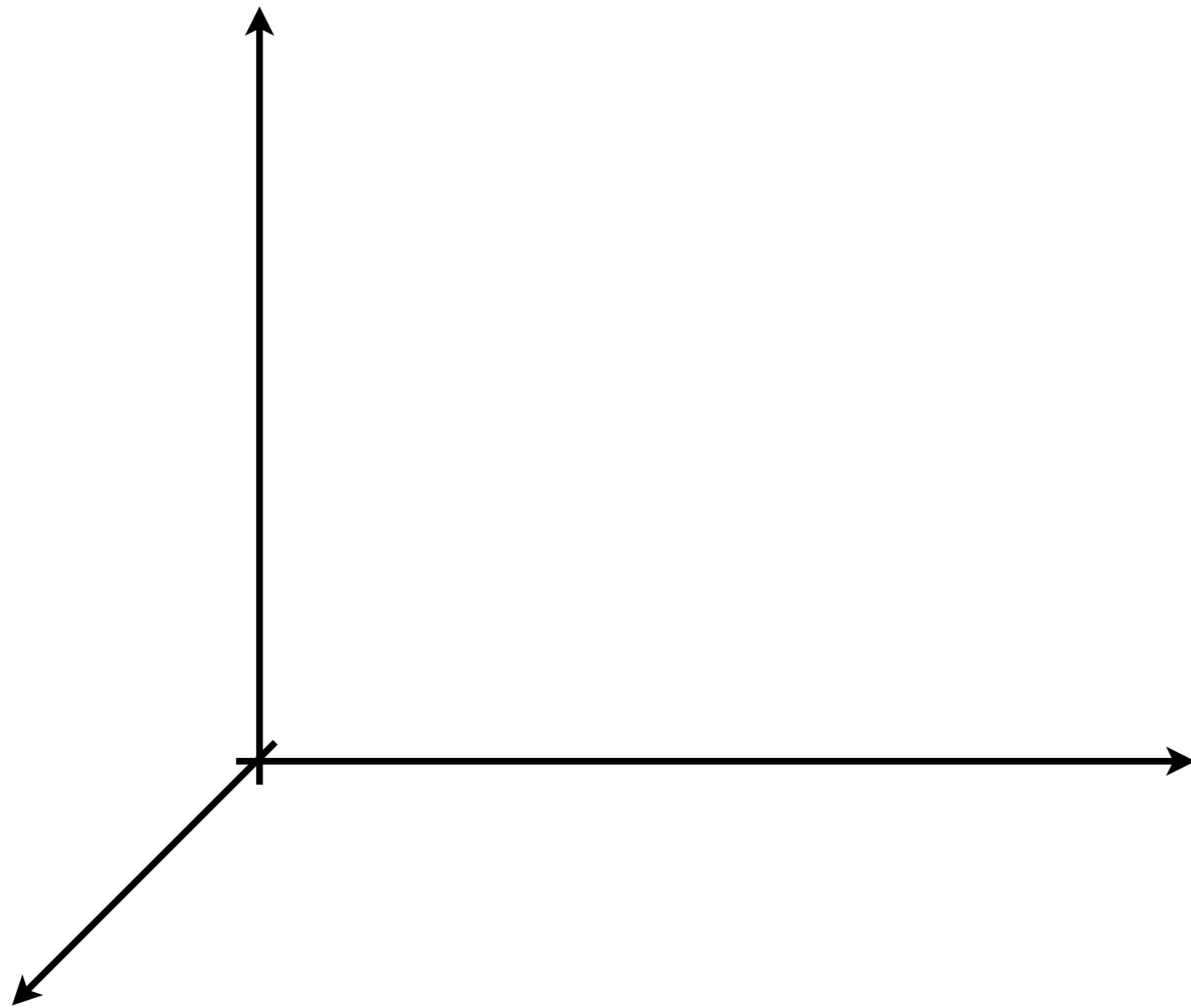
Feature
extraction



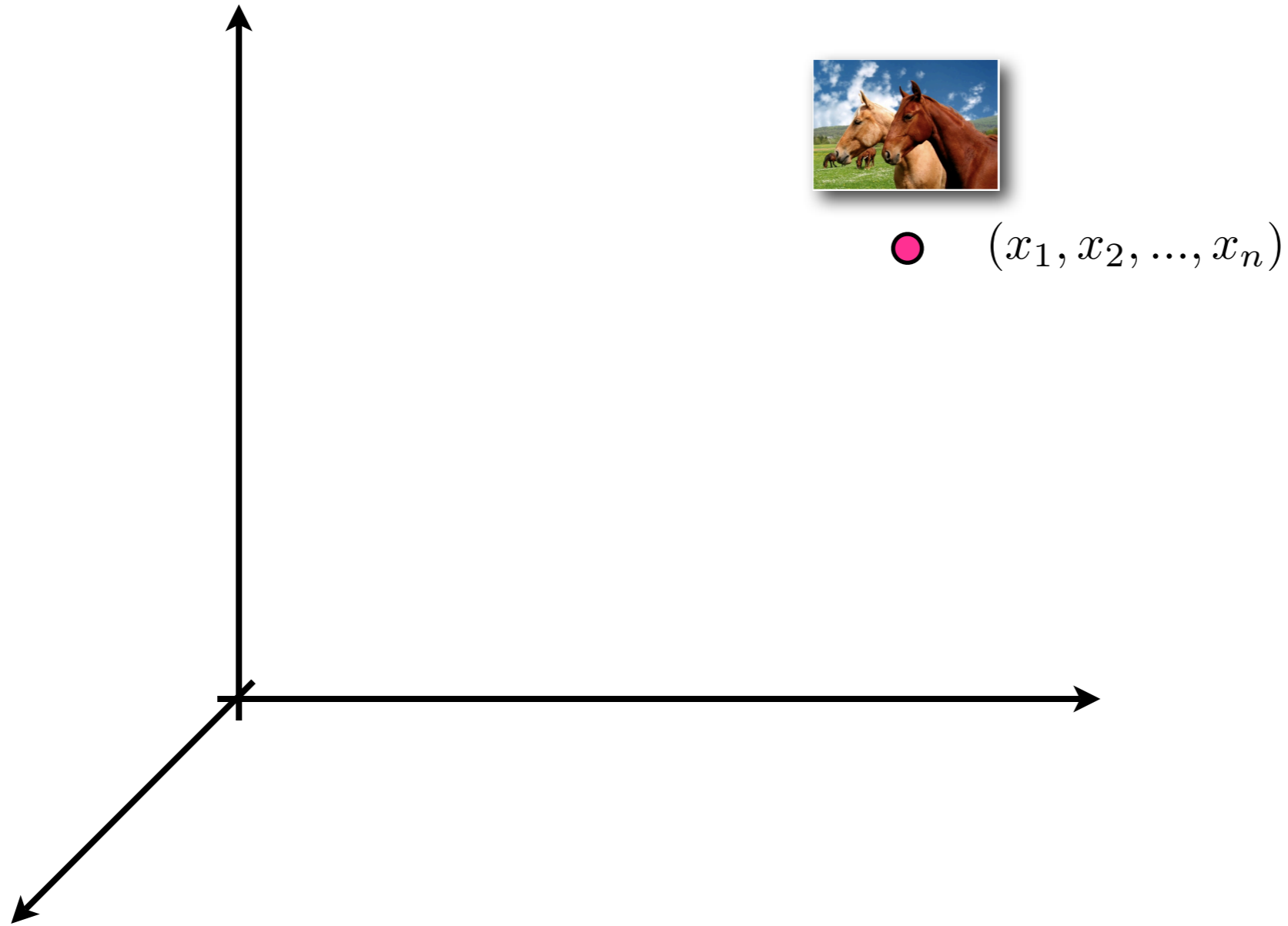
Image B

$$\longrightarrow (y_1, y_2, \dots, y_n)$$

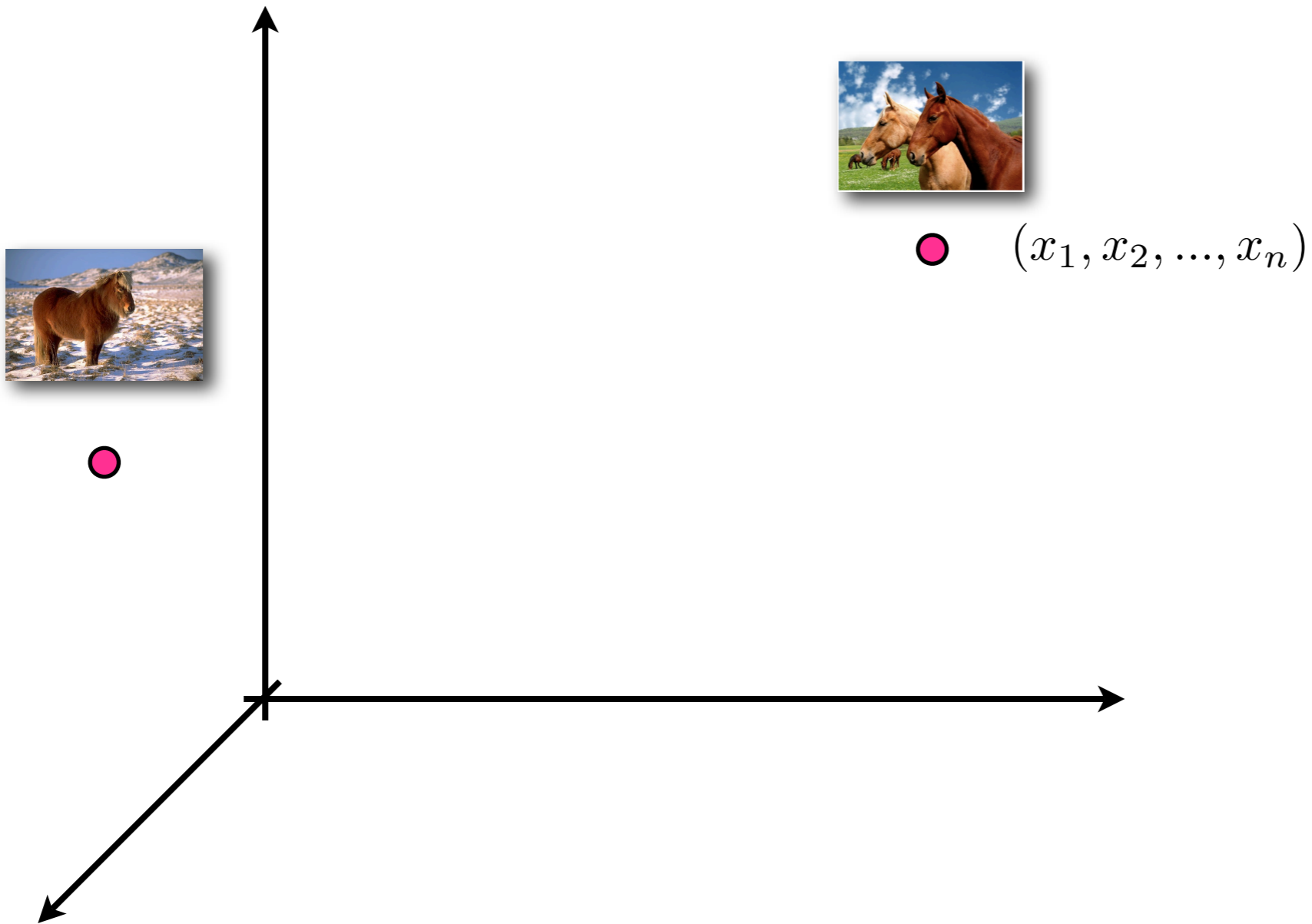
A geometrical view of CBIR



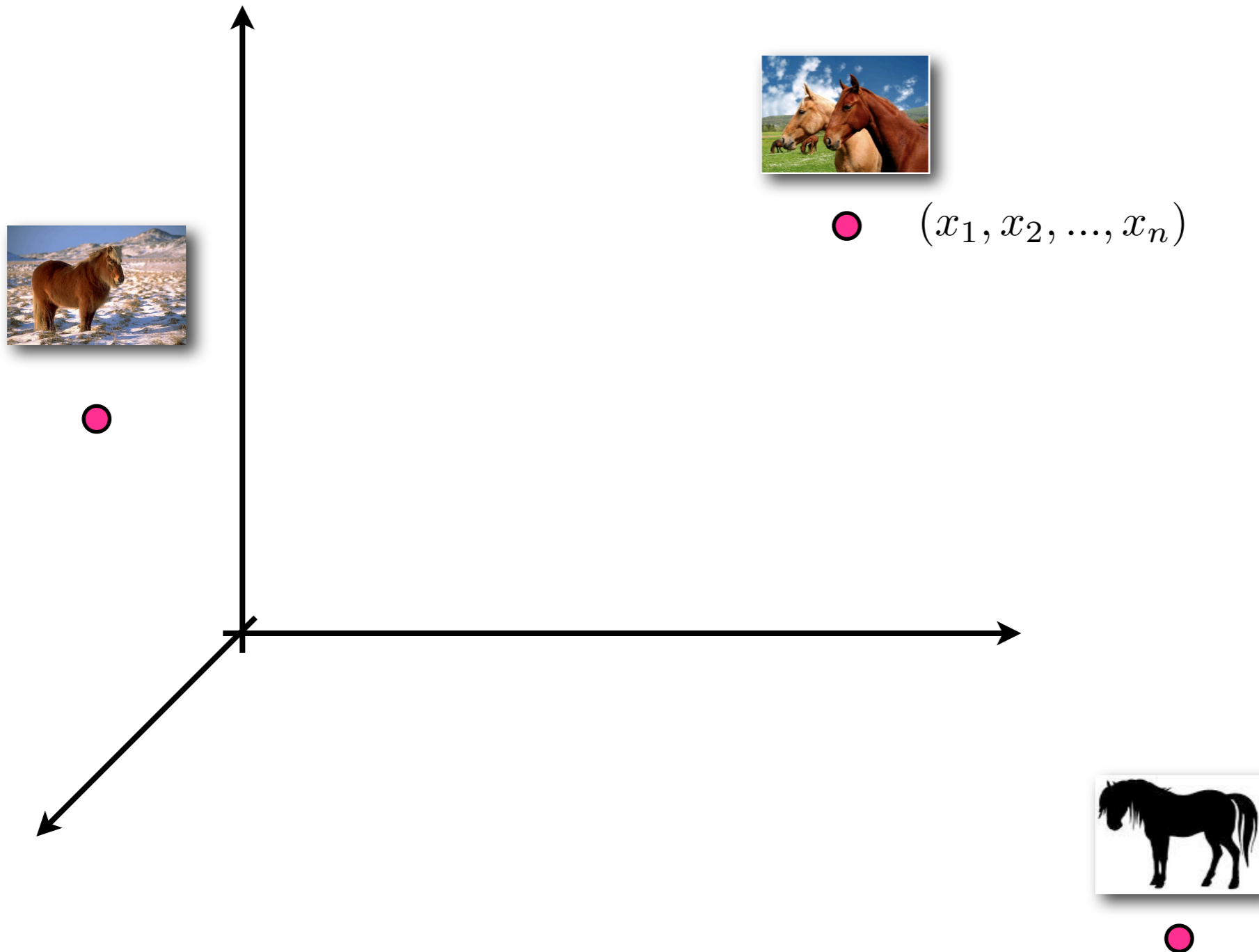
A geometrical view of CBIR



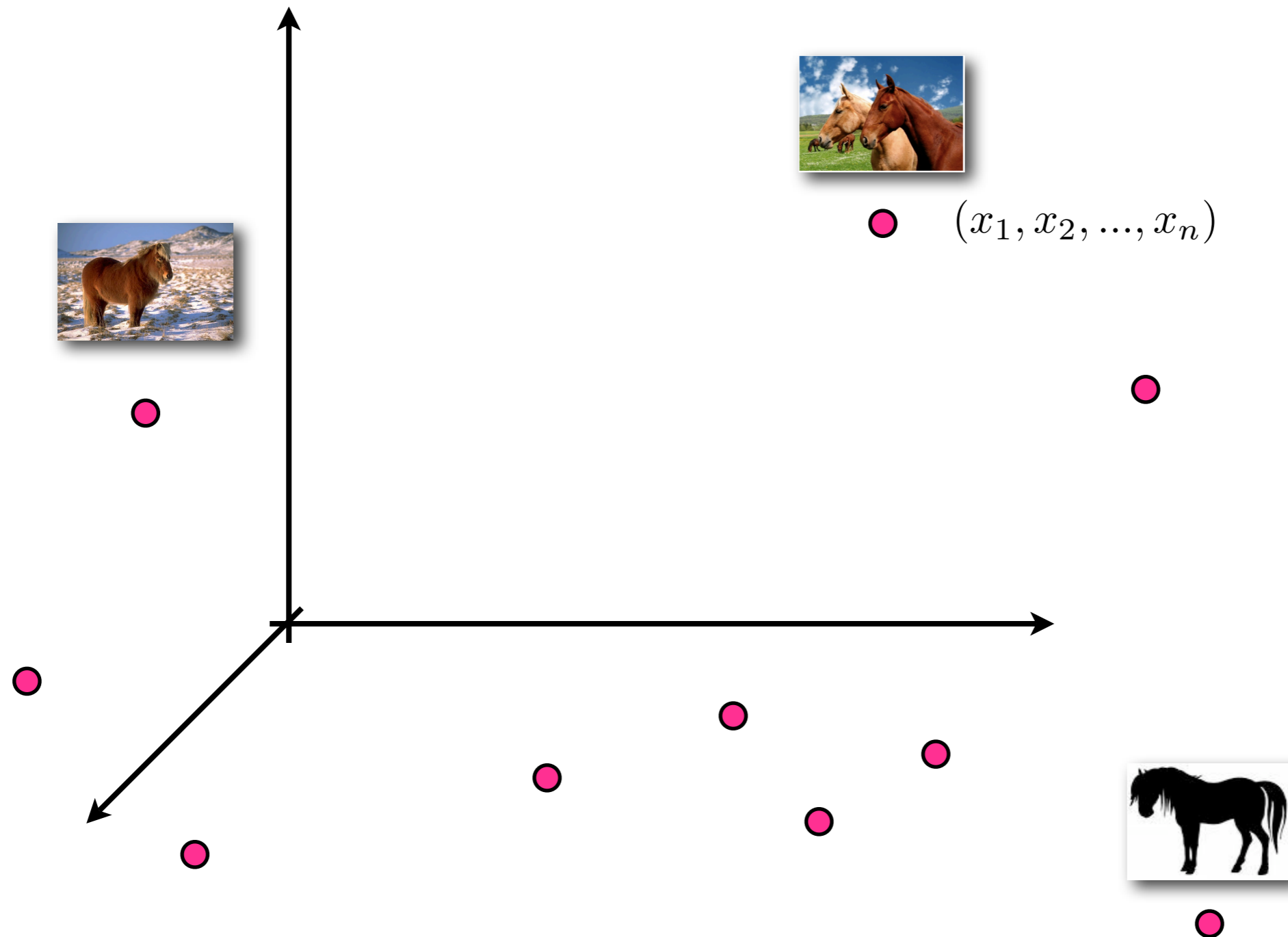
A geometrical view of CBIR



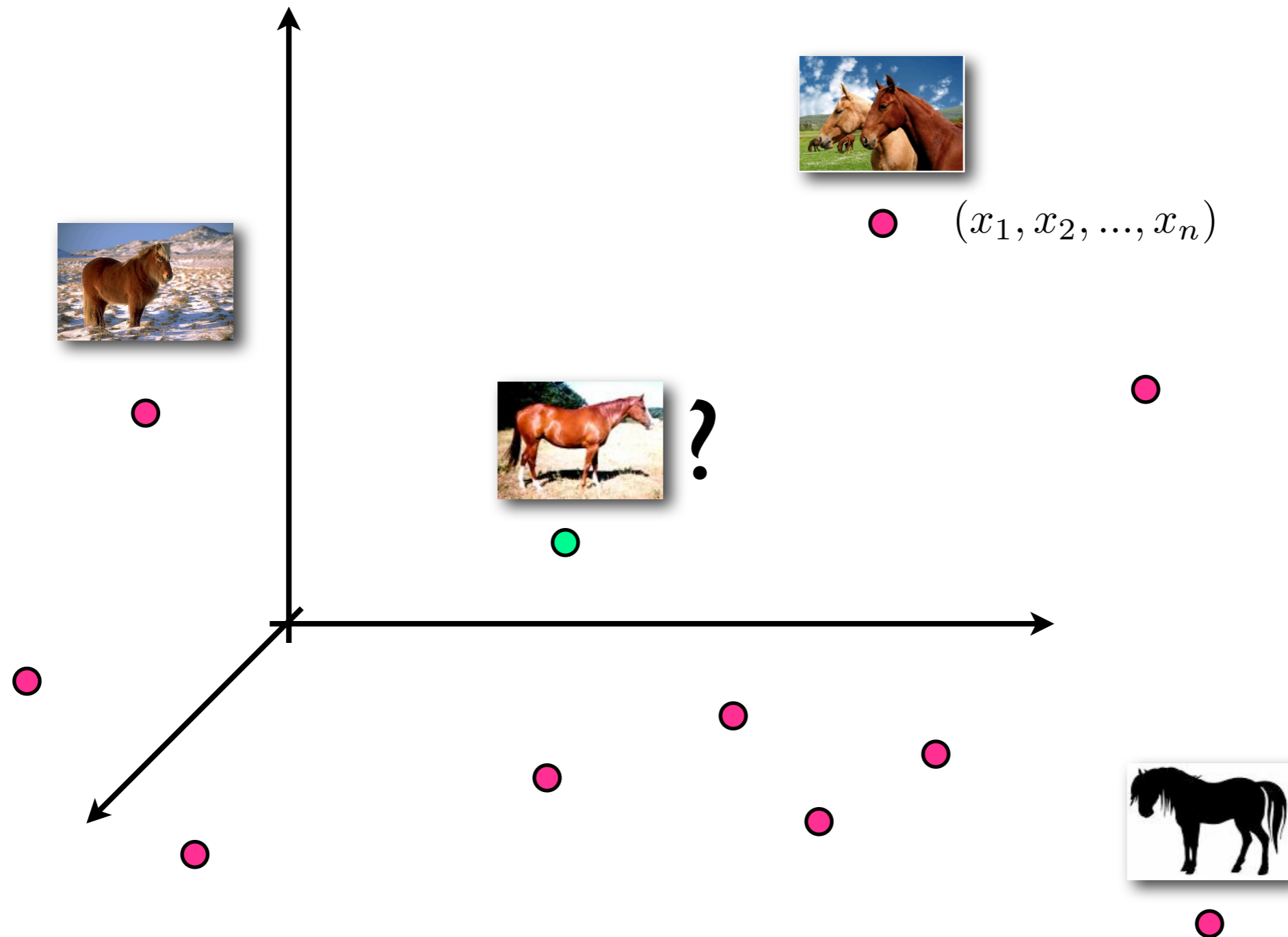
A geometrical view of CBIR



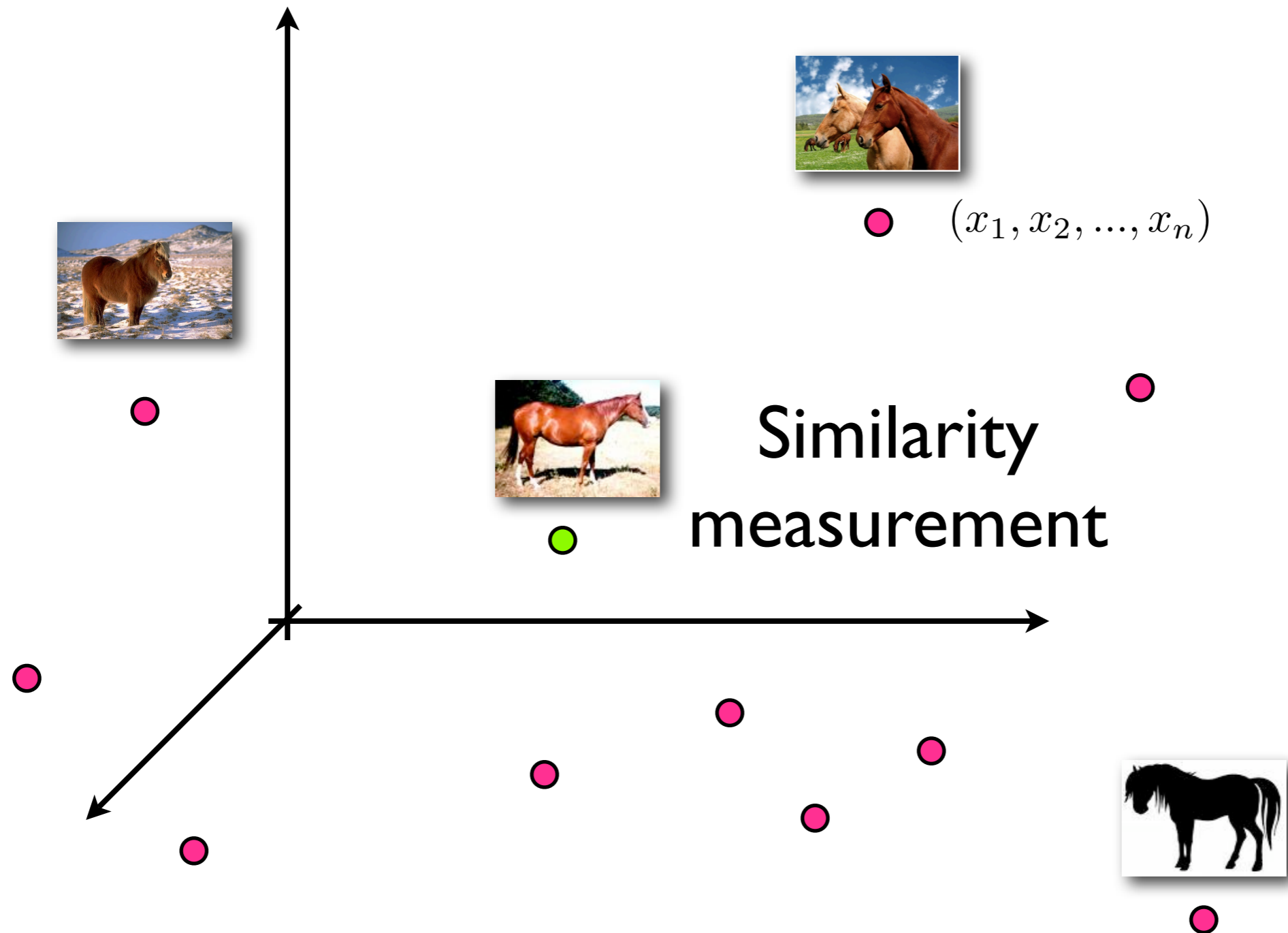
A geometrical view of CBIR



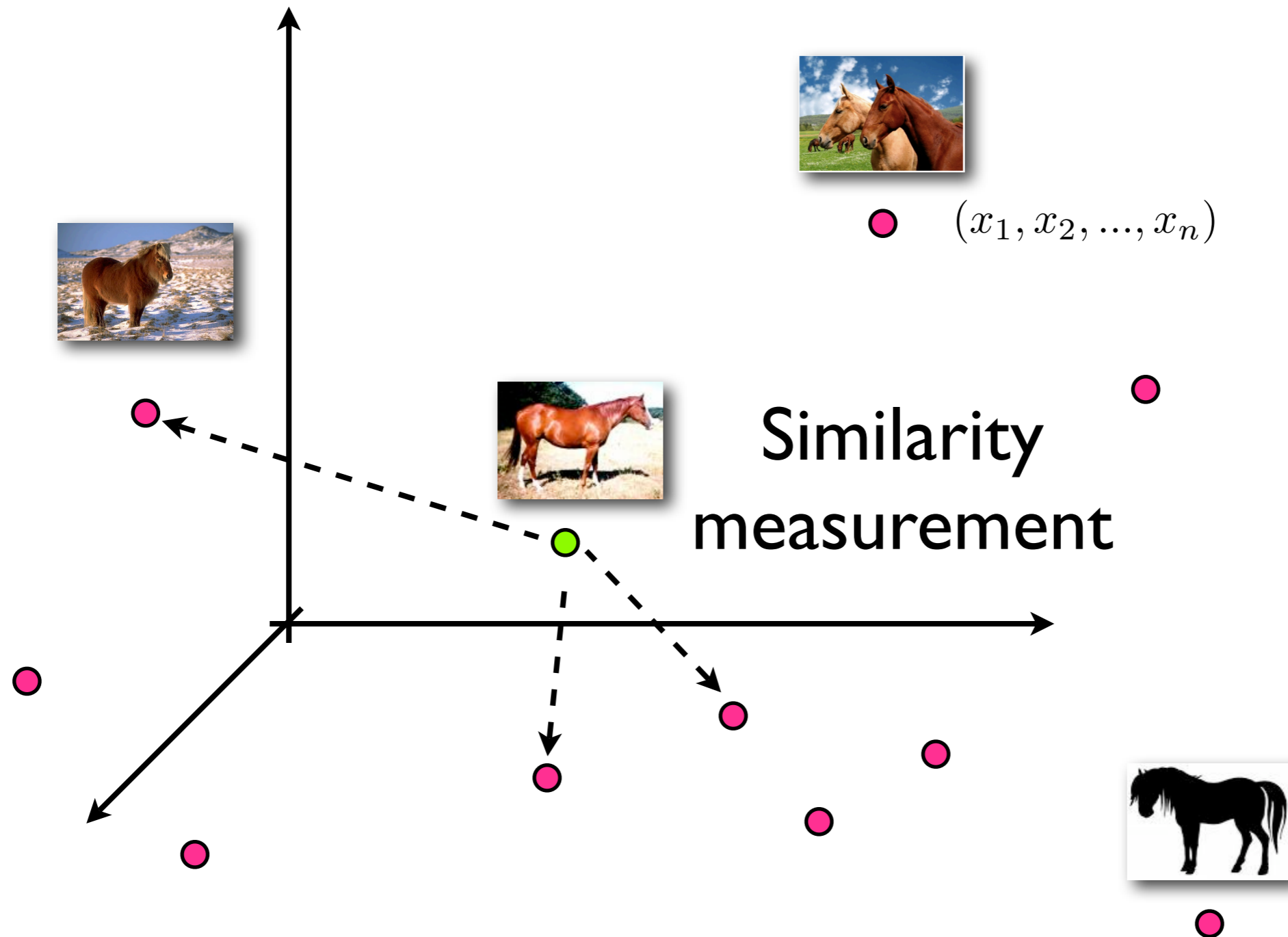
A geometrical view of CBIR



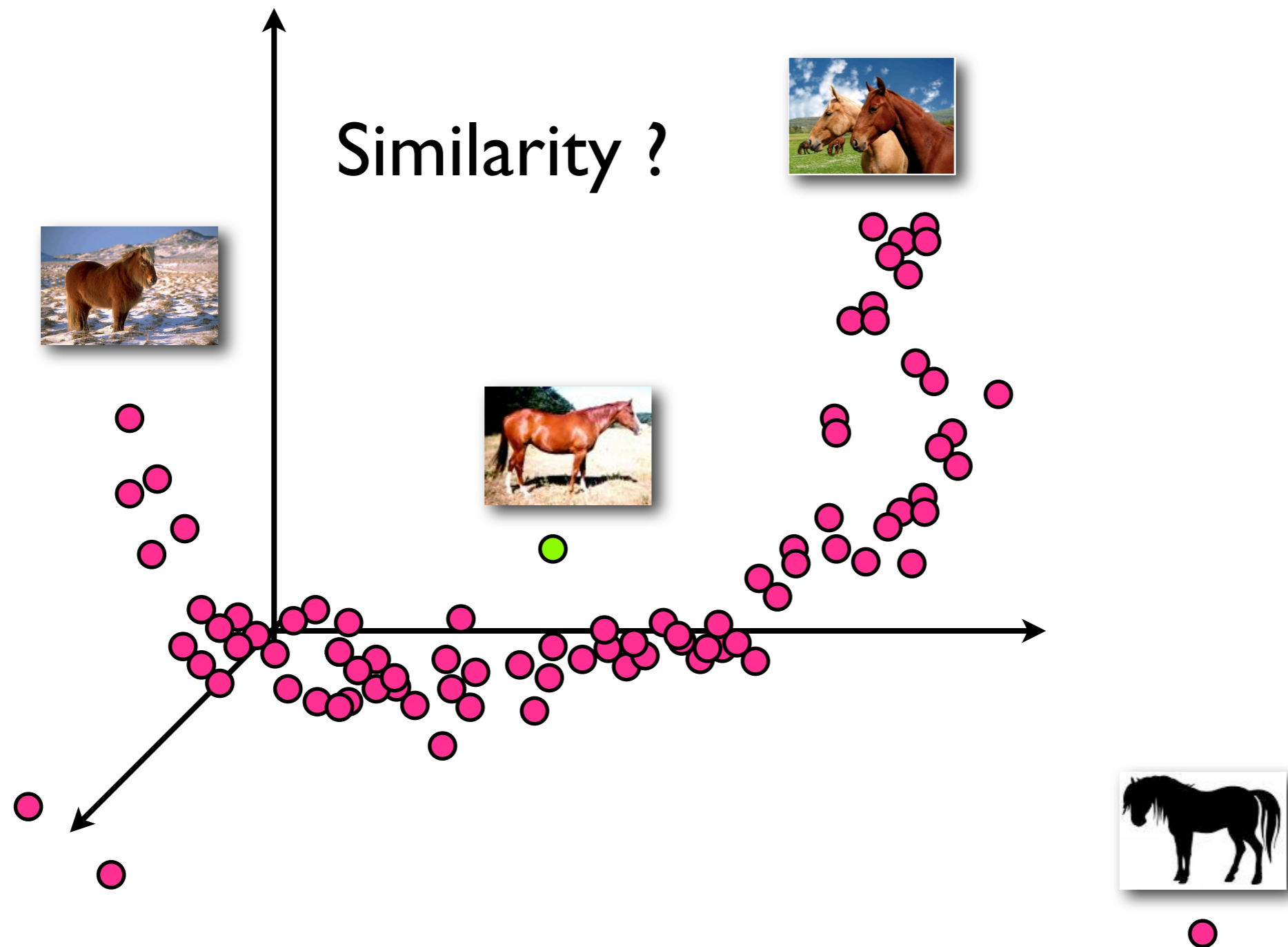
A geometrical view of CBIR



A geometrical view of CBIR



A geometrical view of CBIR



A geometrical view of CBIR

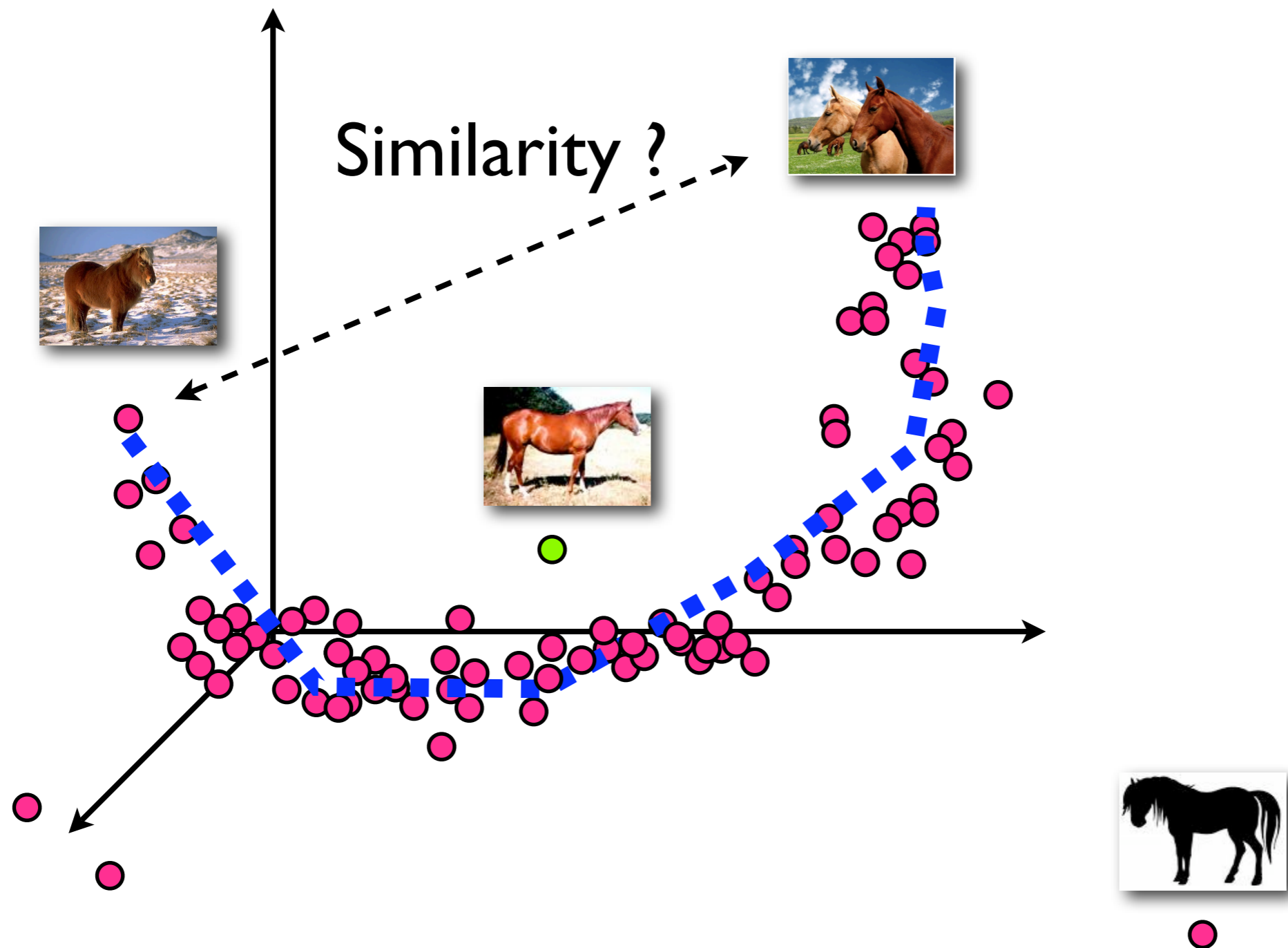


Image similarities

- How to measure similarity of different images based on features?
 - Image features always form into a fixed-length feature vector.
 - The similarity therefore can be measured by
 - Euclidean distance
 - Histogram intersection
 - Quadratic distance
 - Mahalanobis distance (马氏距离)
 - Non-geometrical similarity



Practical image retrieval systems

- QBIC (Query By Image Content)
 - <http://www.qbic.almaden.ibm.com/>
- Virage
 - <http://www.virage.com/cgi-bin/query-e>
- RetrievalWare
 - <http://vrw.excalib.com/cgi-bin/sdk/cst/cst2.bat>
- Photobook
- MARS
 - <http://jadzia.ifp.uiuc.edu:8000>

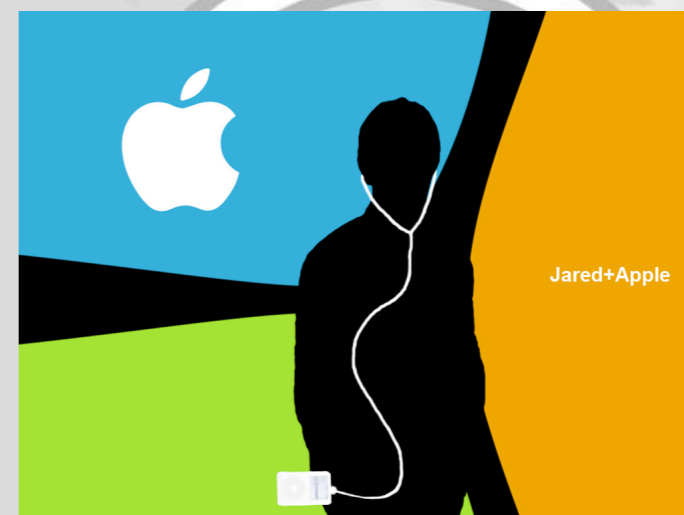
Practical image retrieval systems (cont.)

- Most existing image retrieval systems have one or more of following functions features:
 - Random browsing
 - Classified browsing
 - Example based retrieval
 - Sketch based retrieval
 - Texture based retrieval





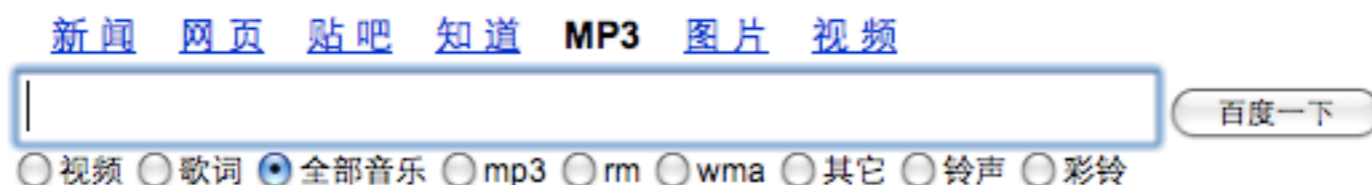
2. music retrieval techniques



reference:

Content based music retrieval

- 說明
 - 用聲音的內容為根據，做音樂的檢索
- 目的
 - 讓使用者可以用自然的方式點選歌曲



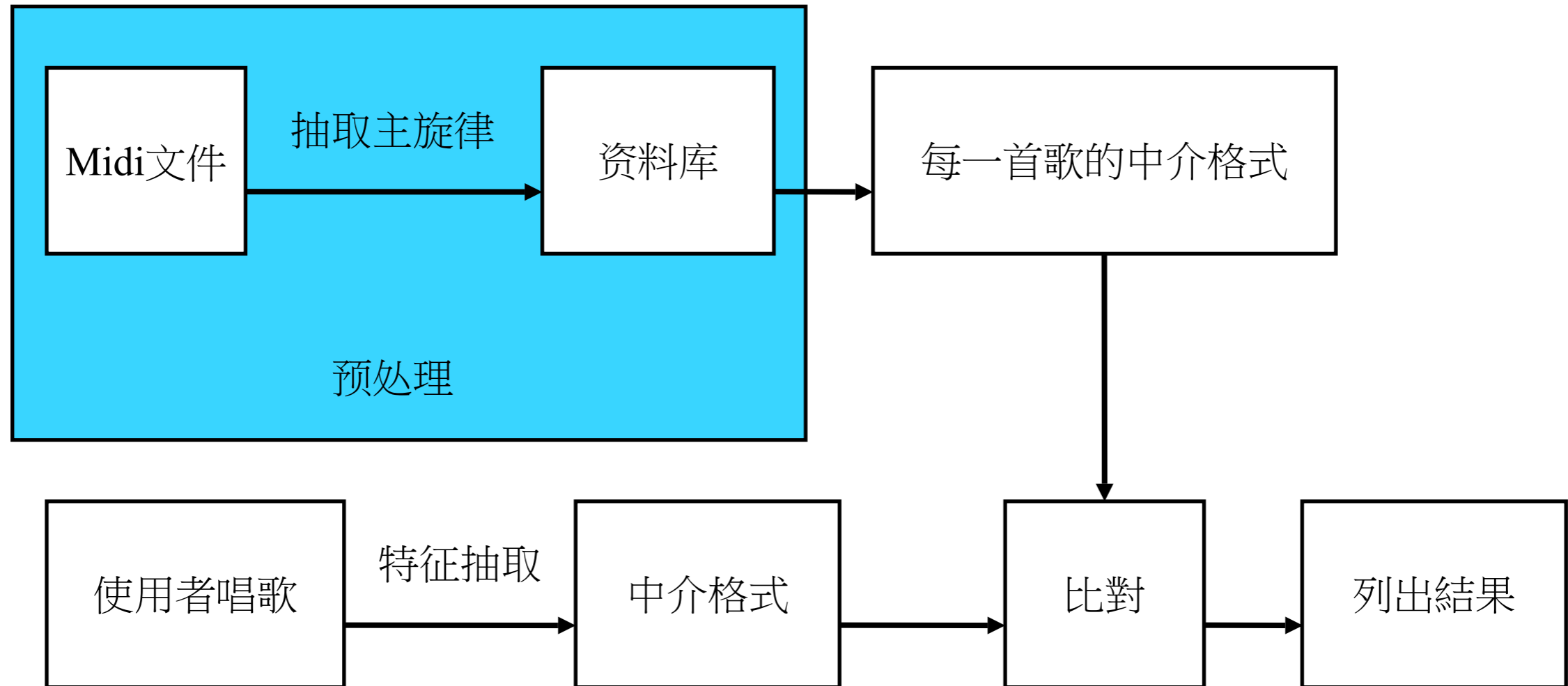
Content based music retrieval

- 說明
 - 用聲音的內容為根據，做音樂的檢索
- 目的
 - 讓使用者可以用自然的方式點選歌曲

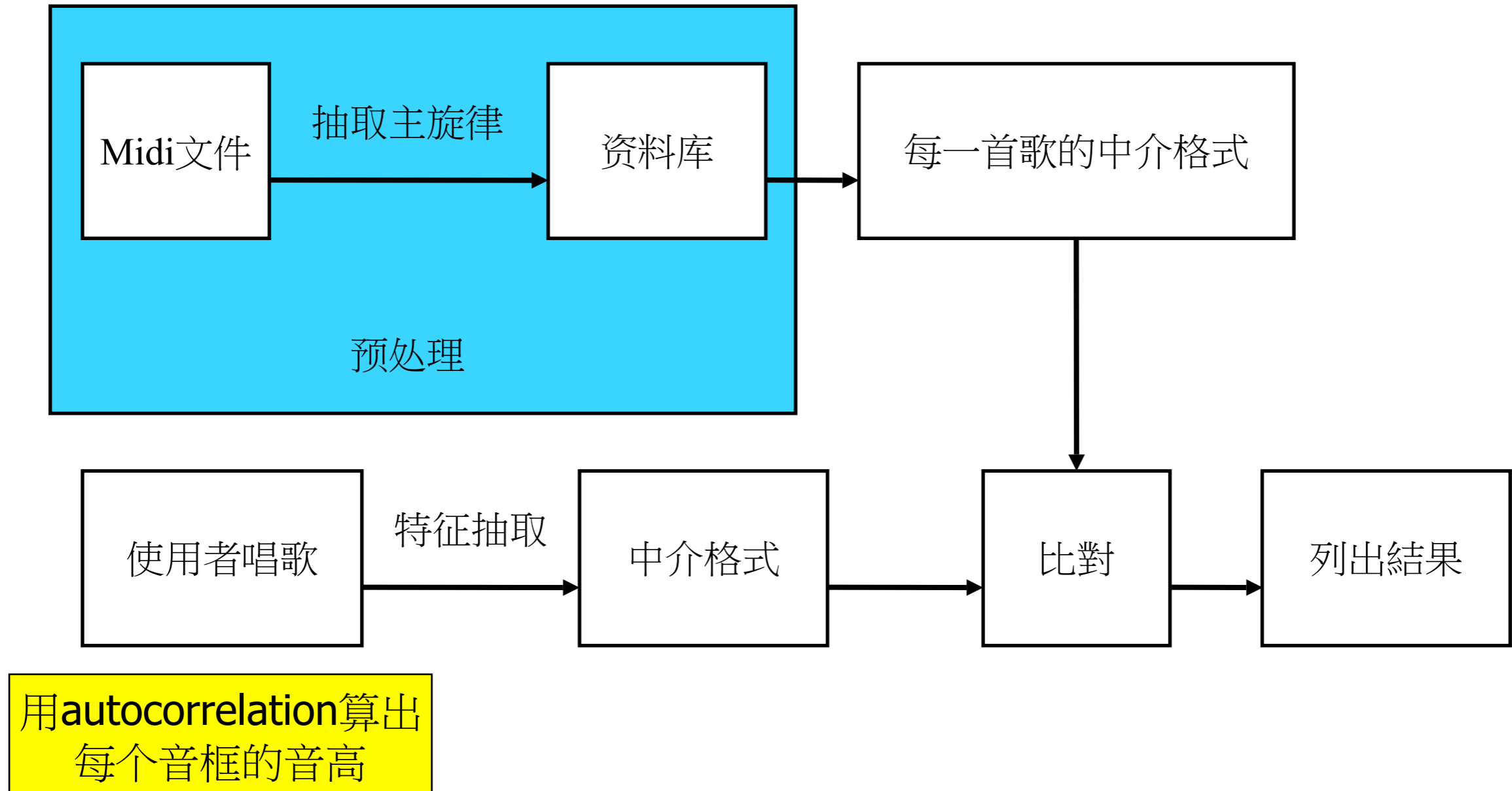
Content based music retrieval

- 困難
 - 使用者的節奏（**tempo**）快慢不同、拍子不準、音調（**key**）高低不同
 - 若允許使用者從歌的任意處唱，計算量很大

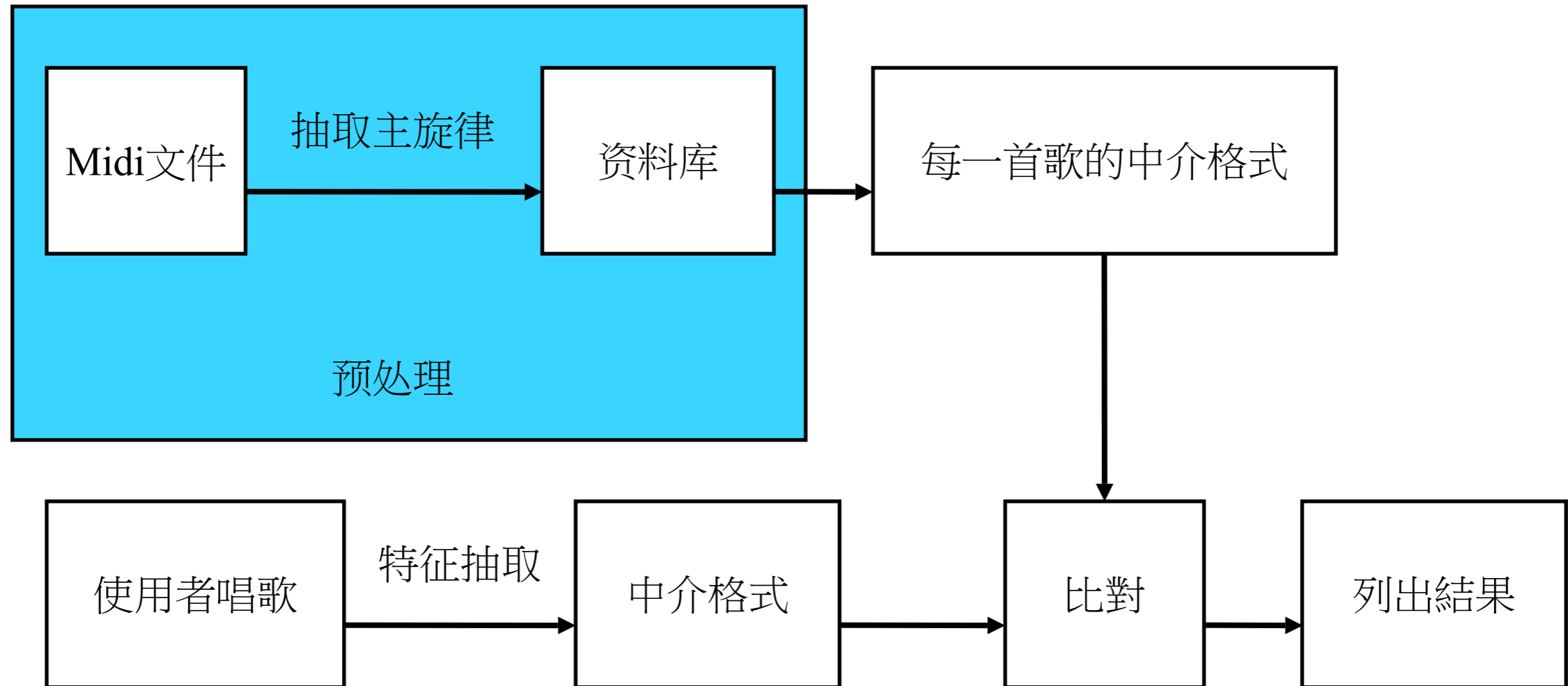
CBMR系统流程图



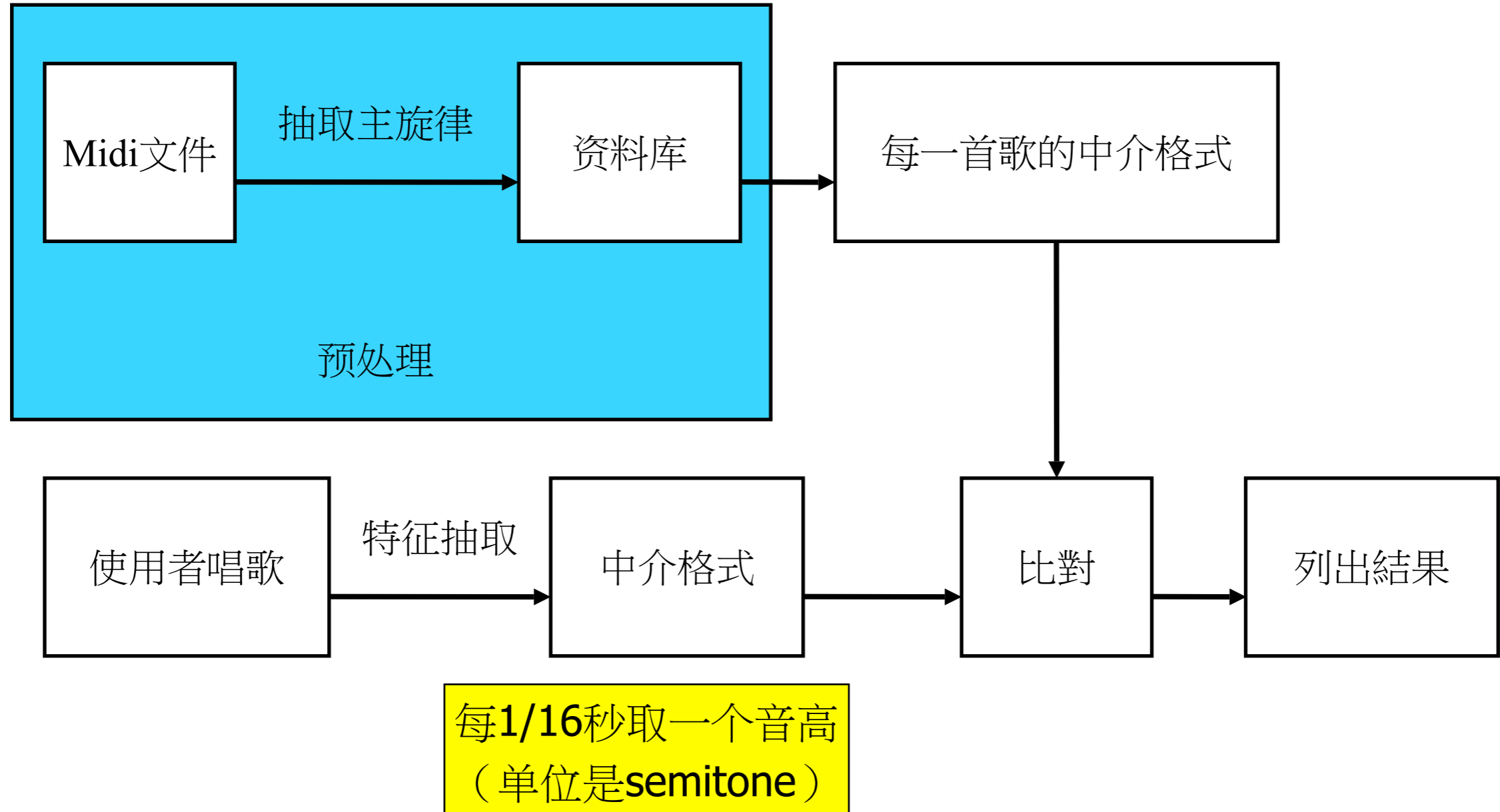
CBMR系统流程图



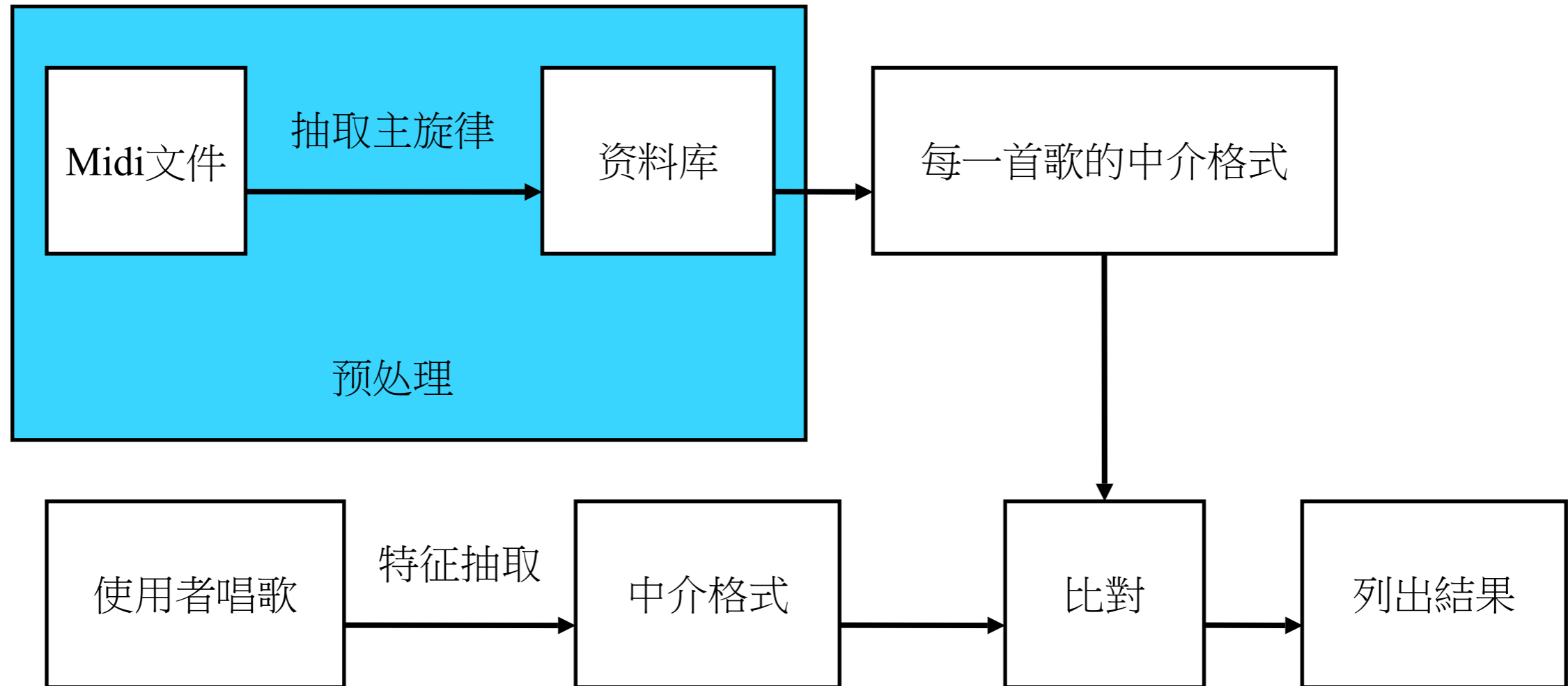
CBMR系统流程图



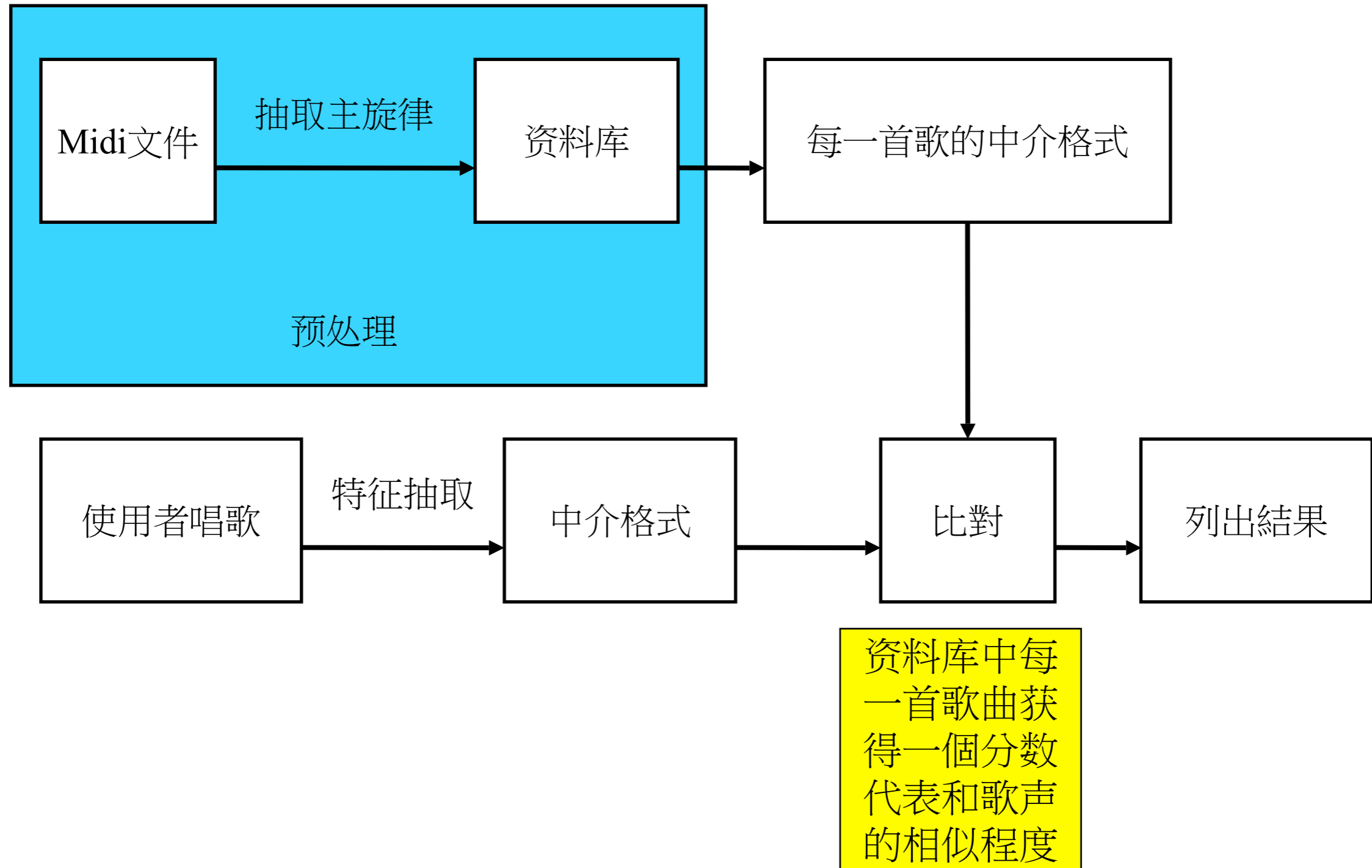
CBMR系统流程图



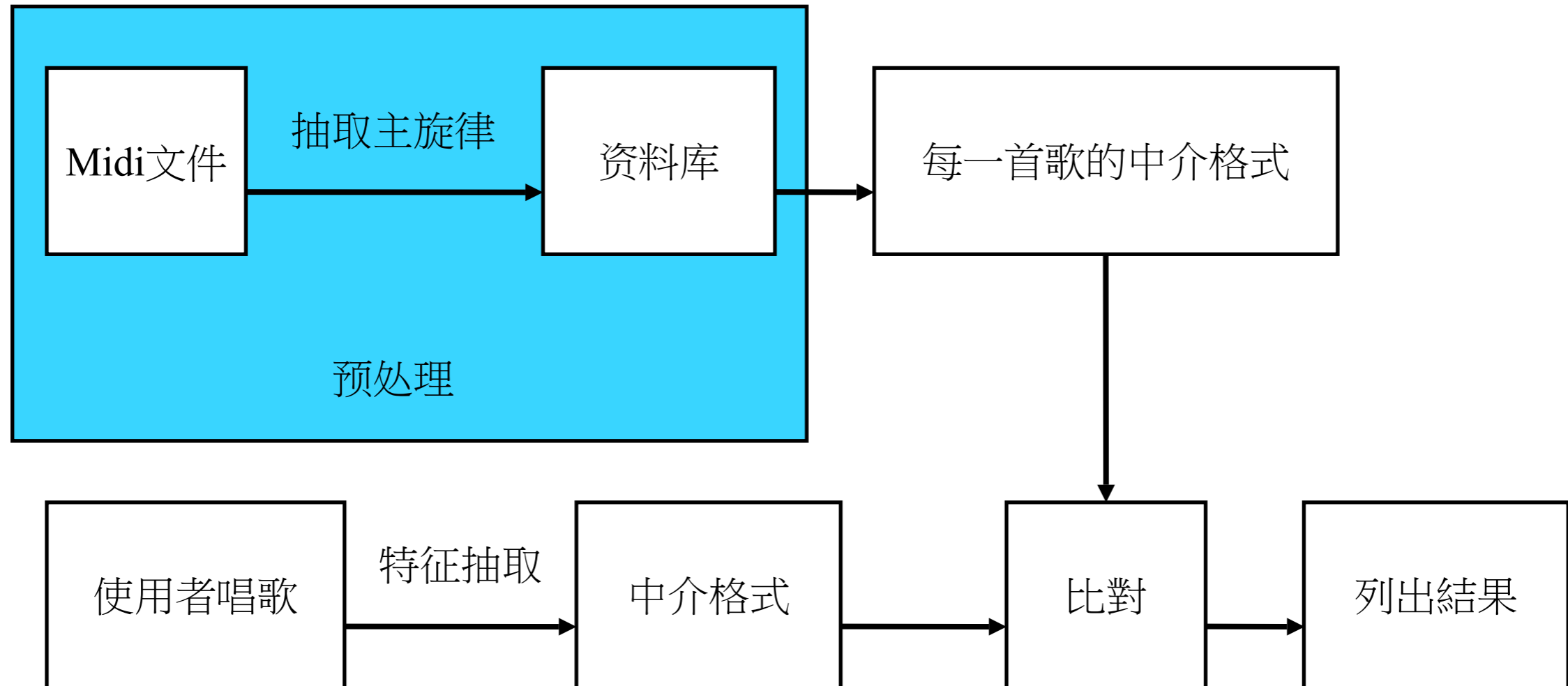
CBMR系统流程图



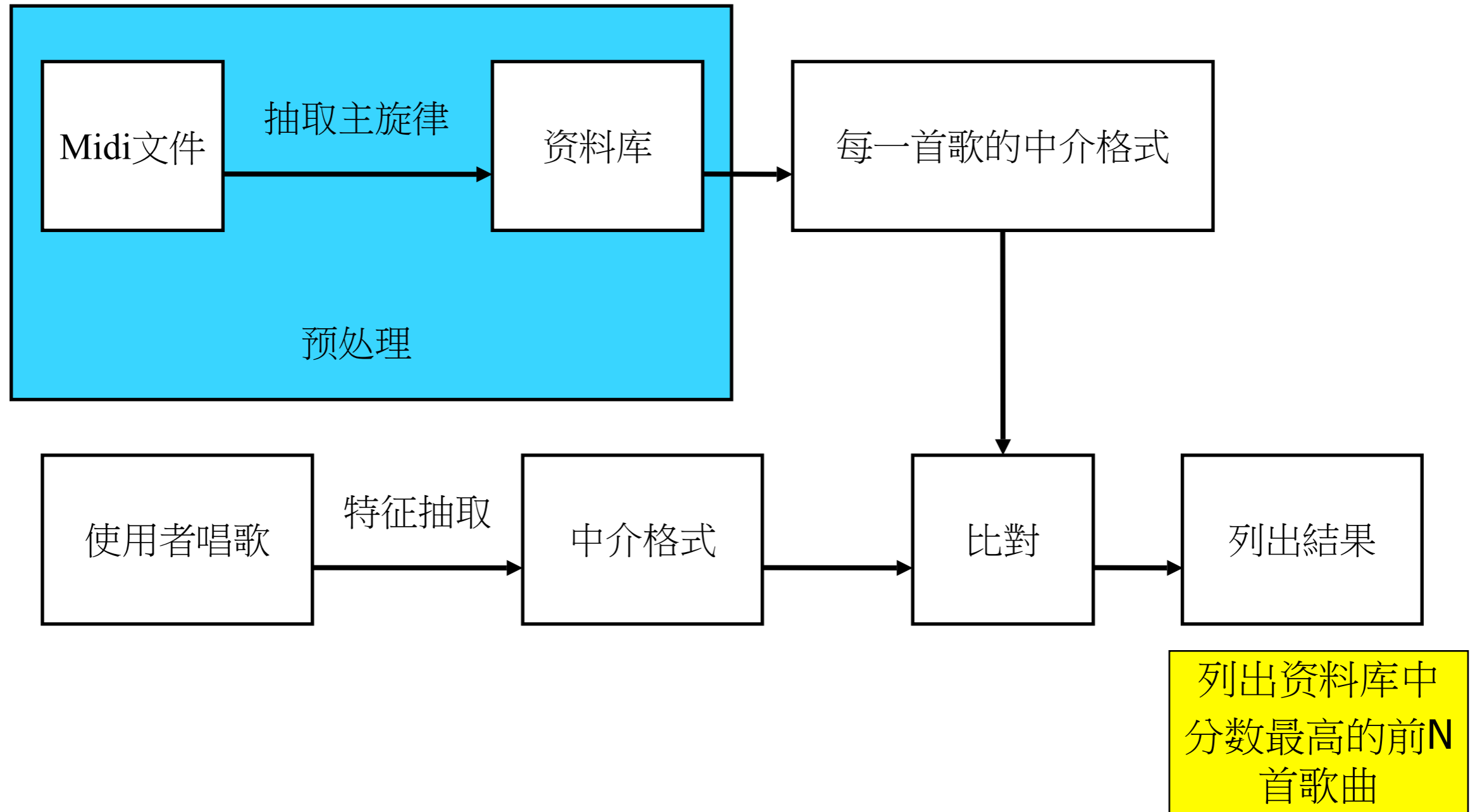
CBMR系统流程图



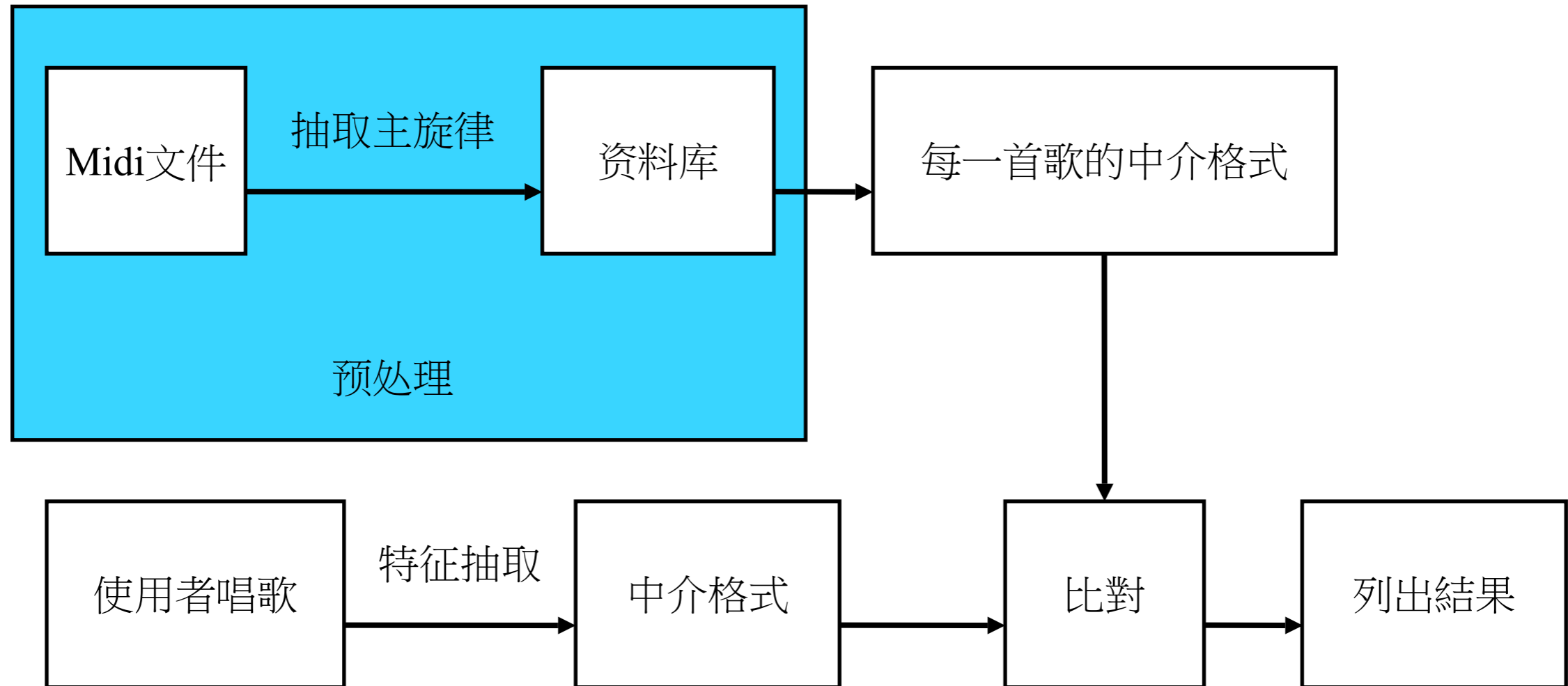
CBMR系统流程图



CBMR系统流程图

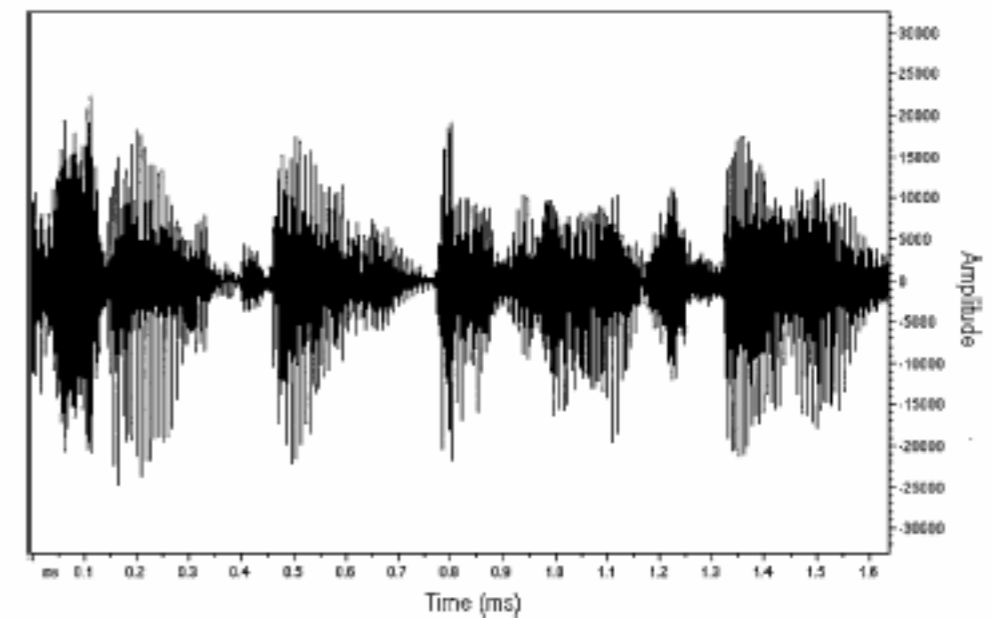


CBMR系统流程图



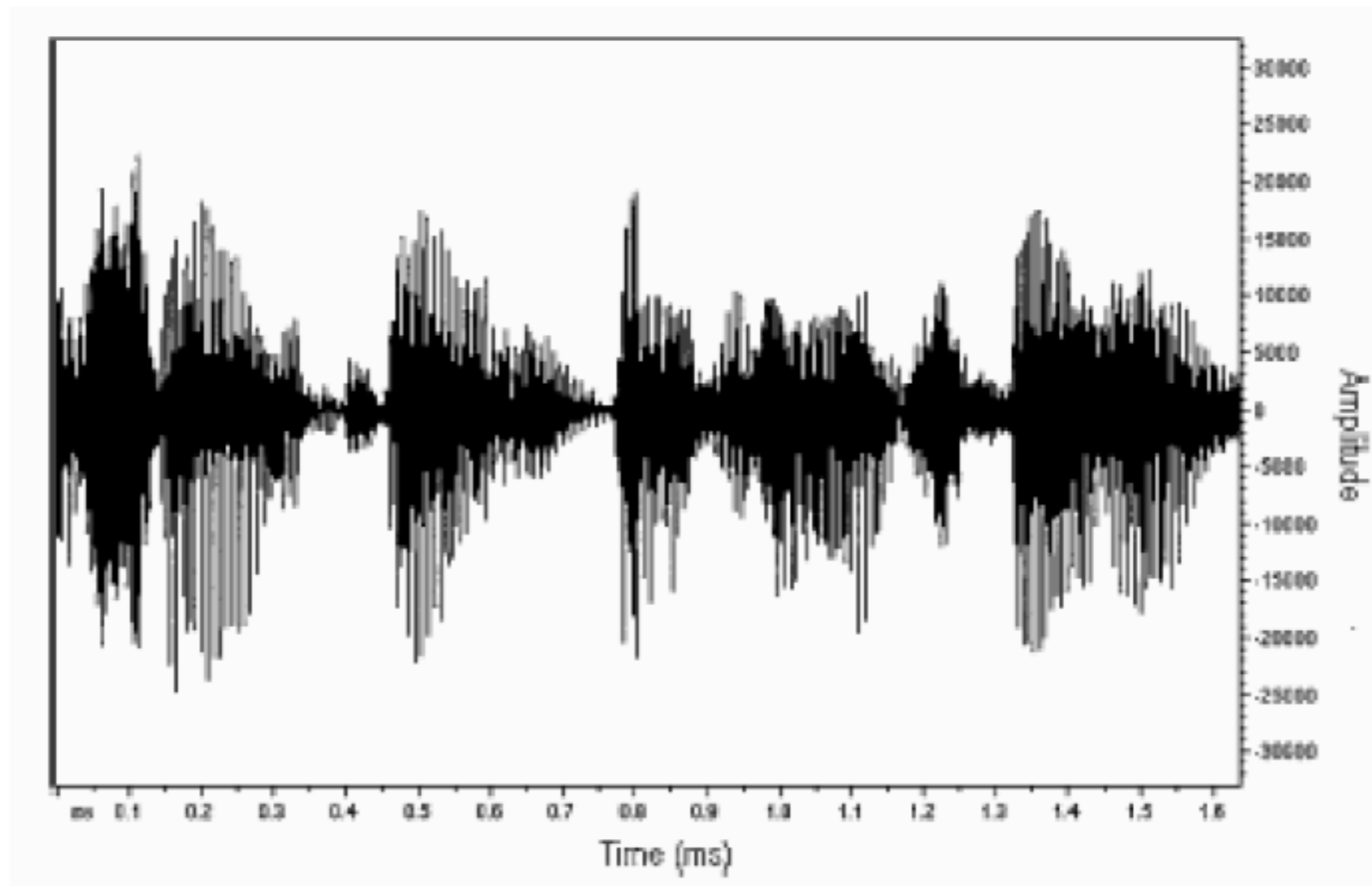
Main Audio Features

- Time-Domain Features
 - Average Energy
 - Zero Crossing Rate
 - Silence Ratio
- Frequency-Domain Features
 - Sound Spectrum
 - Bandwidth
 - Energy Distribution
 - Harmonicity
 - Pitch
- Spectrogram



Time-Domain Features

- Amplitude-time representation of an audio signal



Time-Domain Features (2)

- Average Energy

- Indicates the loudness of the audio signal

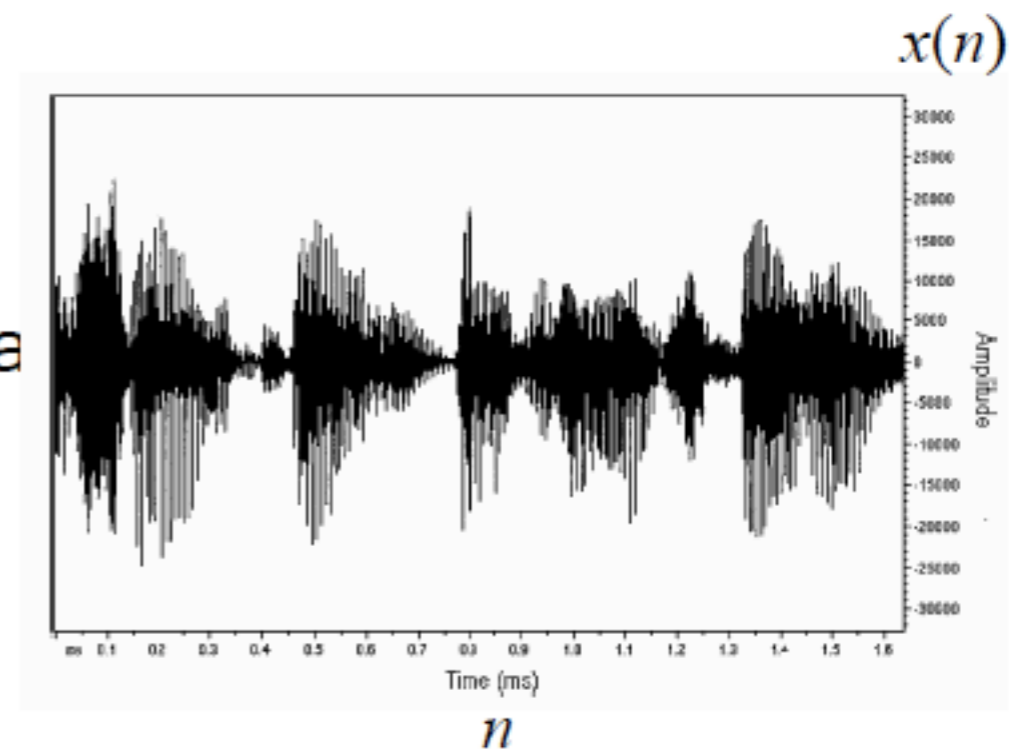
$$E = \frac{\sum_{n=1}^{N-1} x(n)^2}{N}$$

- Zero Crossing Rate

- Indicates the frequency of signal amplitude sign change

$$ZC = \frac{\sum_{n=1}^{N-1} \text{sgn}[x(n)] - \text{sgn}[x(n-1)]}{2N}$$

$$\text{sgn}(a) = \begin{cases} 1 & a > 0 \\ 0 & a = 0 \\ -1 & a < 0 \end{cases}$$

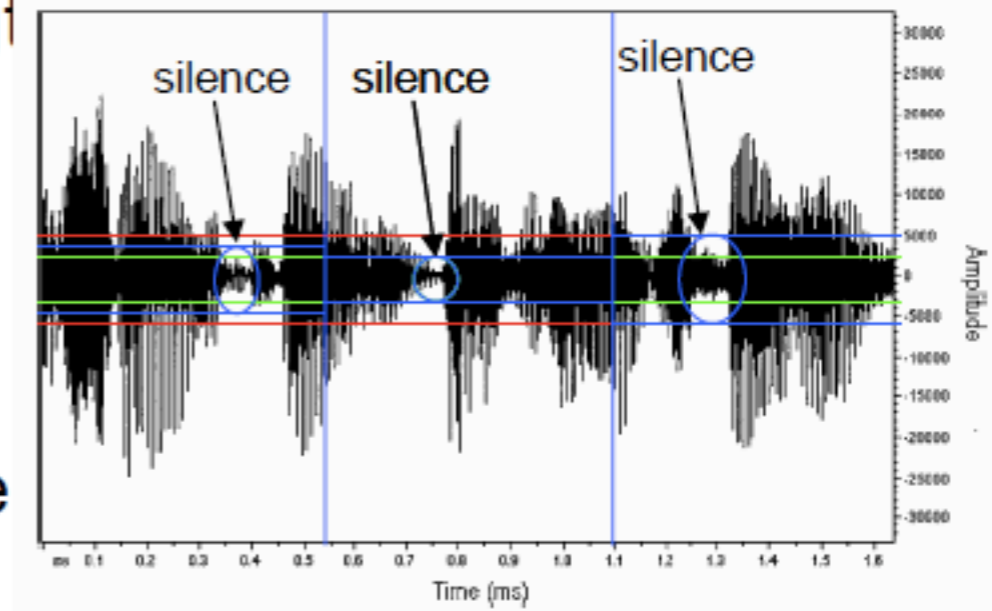


Time-Domain Features (3)

- Silence Ratio
 - Indicates the proportion of the sound piece that is silent
 - Silence is a period within which the absolute amplitude values of a certain number of samples are below a certain threshold
 - Silence ratio is calculated as the ratio between the sum of silent periods and the total length

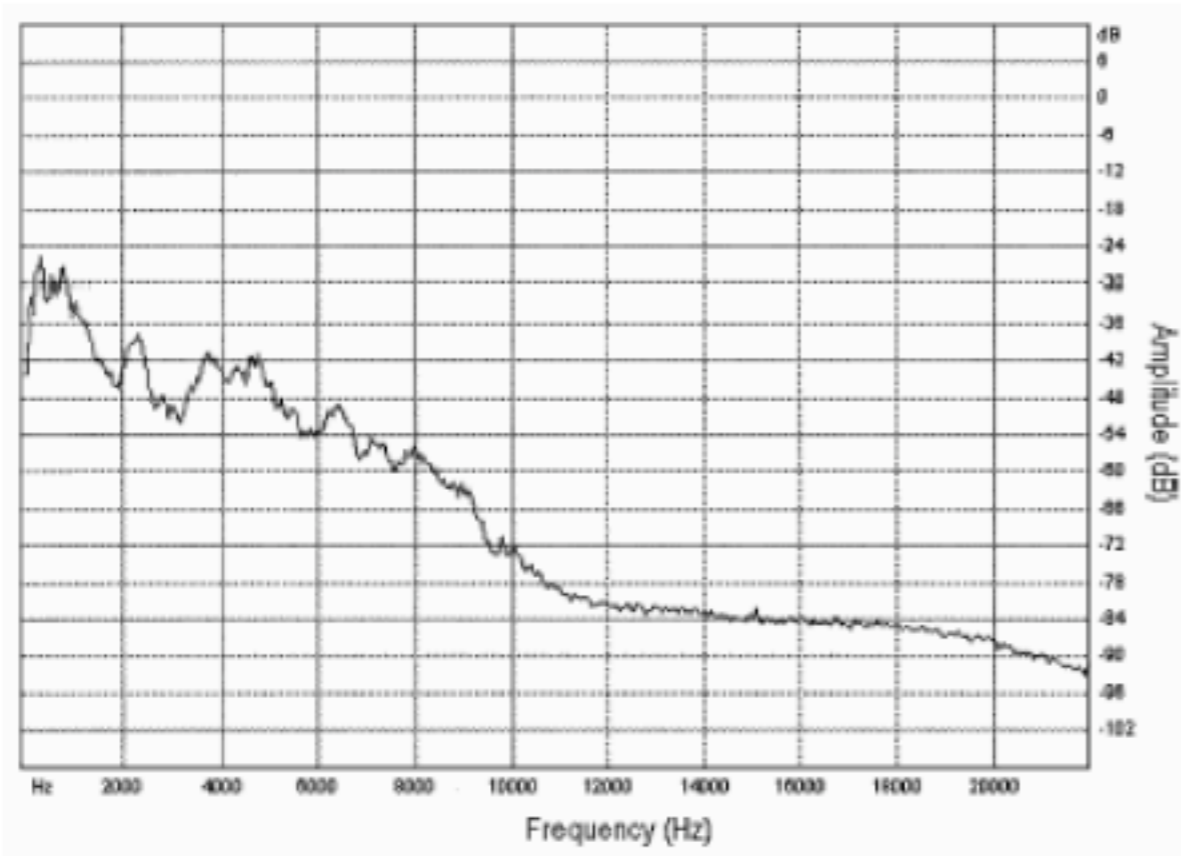
Approaches:

1. Fixed Threshold
2. Select Reference Silence Value
3. Adaptive Silence Thresholds



Frequency-Domain Features

- Sound Spectrum



Discrete Fourier Transform (DFT)

$$X(k) = \sum_{n=0}^{N-1} x(n)e^{-\frac{j2\pi nk}{N}}$$

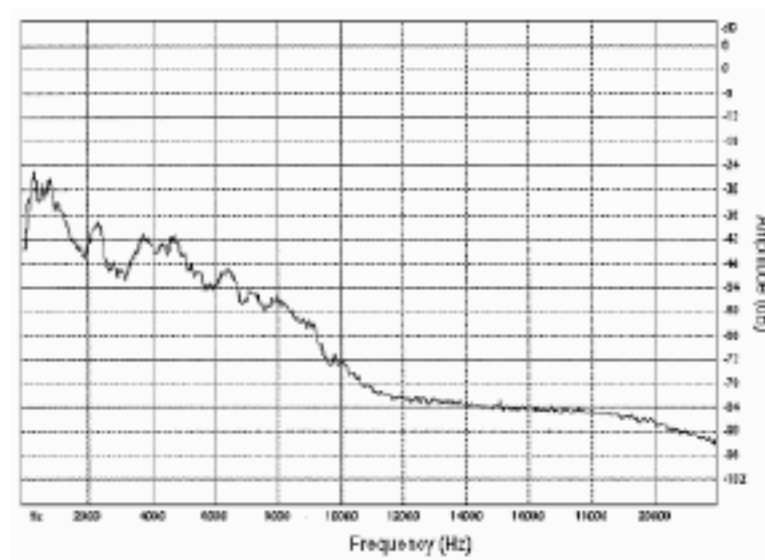
Inverse Discrete Fourier Transform (IDFT)

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{\frac{j2\pi nk}{N}}$$

- For large value of N , the signal is often broken into blocks called frames and DFT is applied to each of the frames.

Frequency-Domain Features (2)

- Bandwidth
 - indicated the frequency range of a sound
 - can be taken as the difference between the highest frequency and lowest frequency of non-zero spectrum components
 - “non-zero” may be defined as at least 3dB above the silence level
- Energy distribution
 - Signal distribution across frequency components
 - One important feature derived from the energy distribution is the *centroid*, which is the mid-point of the spectral energy distribution of a sound. Centroid is also called *brightness*



Frequency-Domain Features (3)

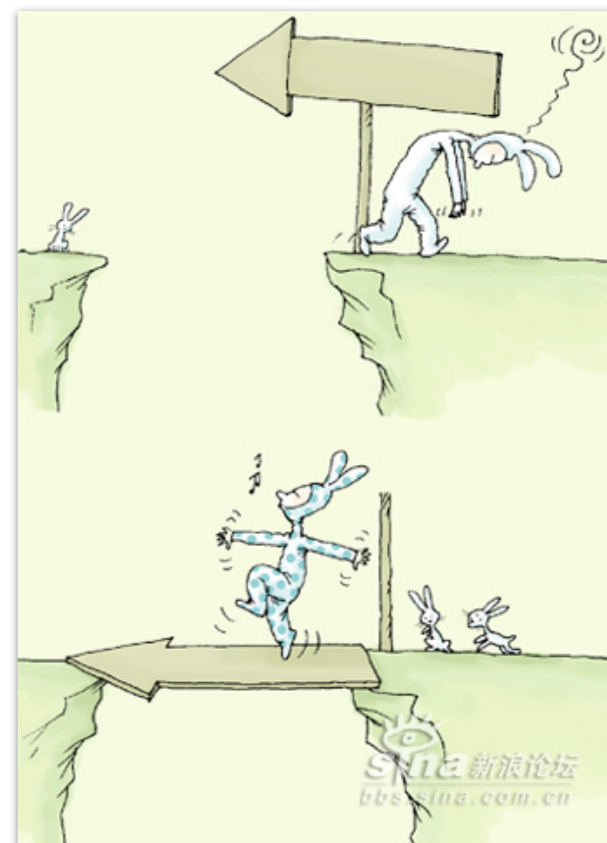
- Harmonicity
 - In harmonic sound, the spectral components are mostly whole number multiples of the lowest and most often loudest frequency
 - Lowest frequency is called *fundamental frequency*
 - Music is normally more harmonic than other sounds
- Pitch
 - the distinctive quality of a sound, dependent primarily on the frequency of the sound waves produced by its source
 - only period sounds, such as those produced by musical instruments and the voice, give rise to a sensation of a pitch
 - In practice, we use the fundamental frequency as the approximation of the pitch

相關研究

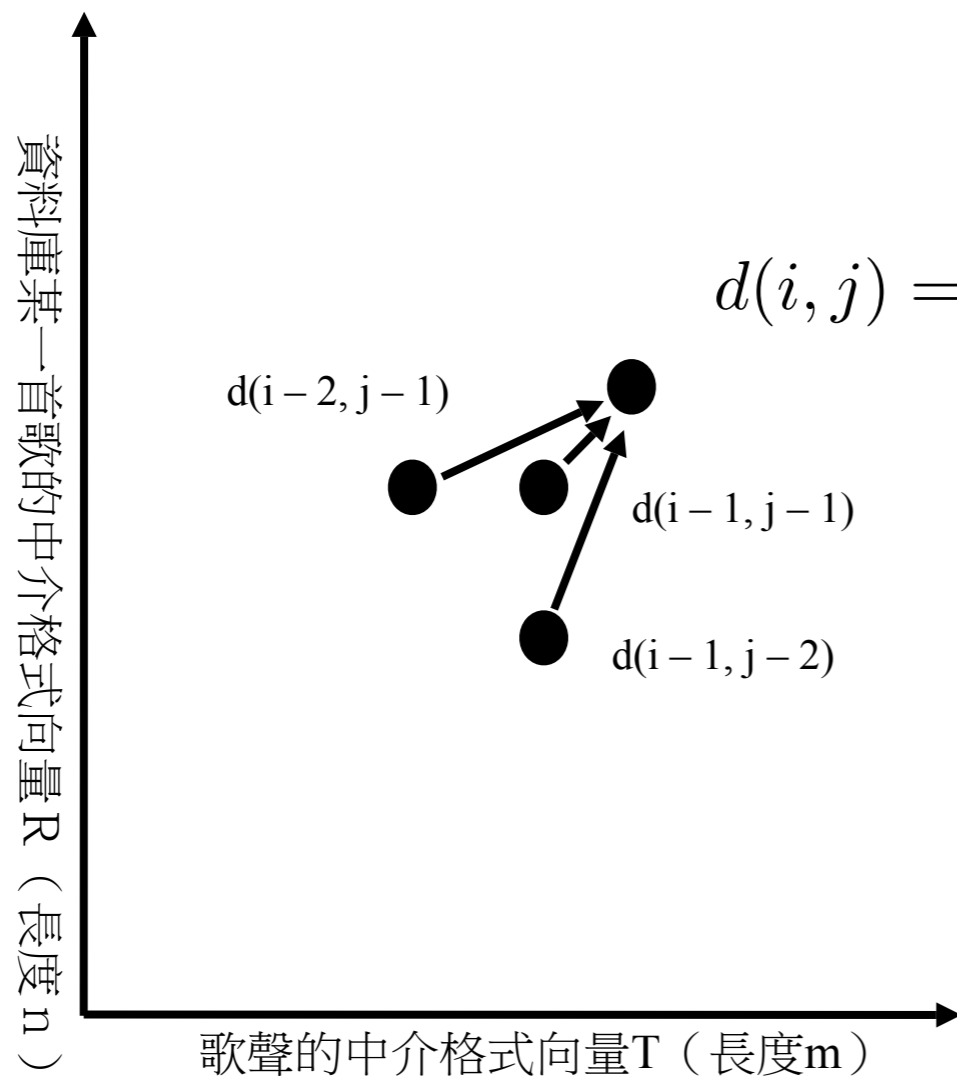
名稱	使用限制	比對方法	資料庫歌曲數目
Query by Humming	必須唱 ta 或 da 斷音	Baeza-Yates & Perleberg (92)	183
MELDEX (Melody Indexing)	必須唱 ta 或 da 斷音	Dynamic programming	9400
SoundCompass	必須照節拍器唱	Pitch transitions & histograms for weighted average	10086
MELODISCOV	無	FIEXPAT, Rolland (99)	未知

前人的方法

- 克服節奏快慢不同的問題
- **Dynamic time warping**
- 克服音調高低不同的問題
- **Key transposition**
 - 將使用者的歌聲和資料庫中的歌轉換後的中介格式的平均值都平移到0，做一次dtw比對
 - 將資料庫中的歌的中介格式上下平移再做四次dtw比對，以找出最短的距離
- 全曲比對費時很久且準確率低
- 使用浮點數運算



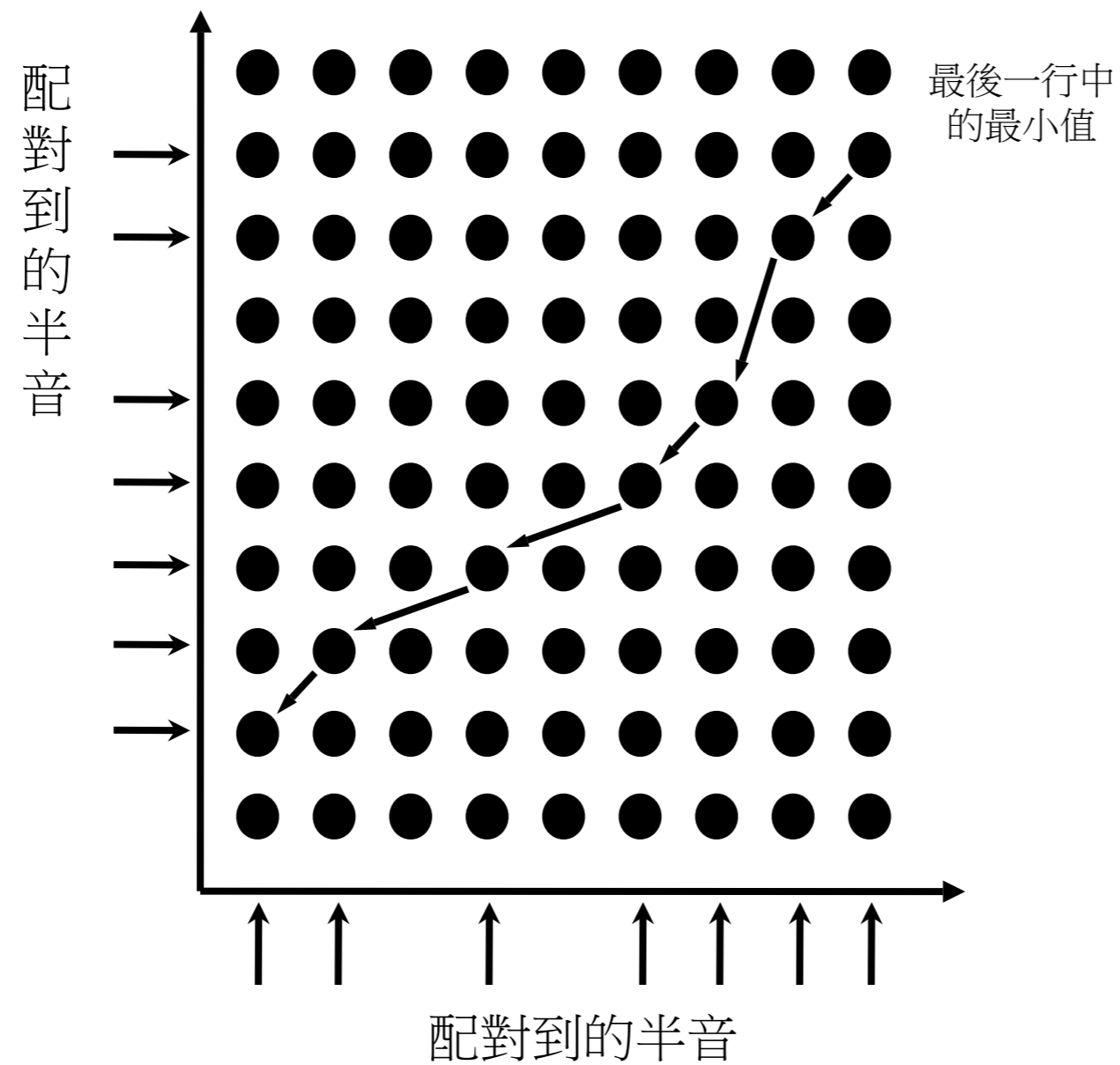
Dynamic Time Warping



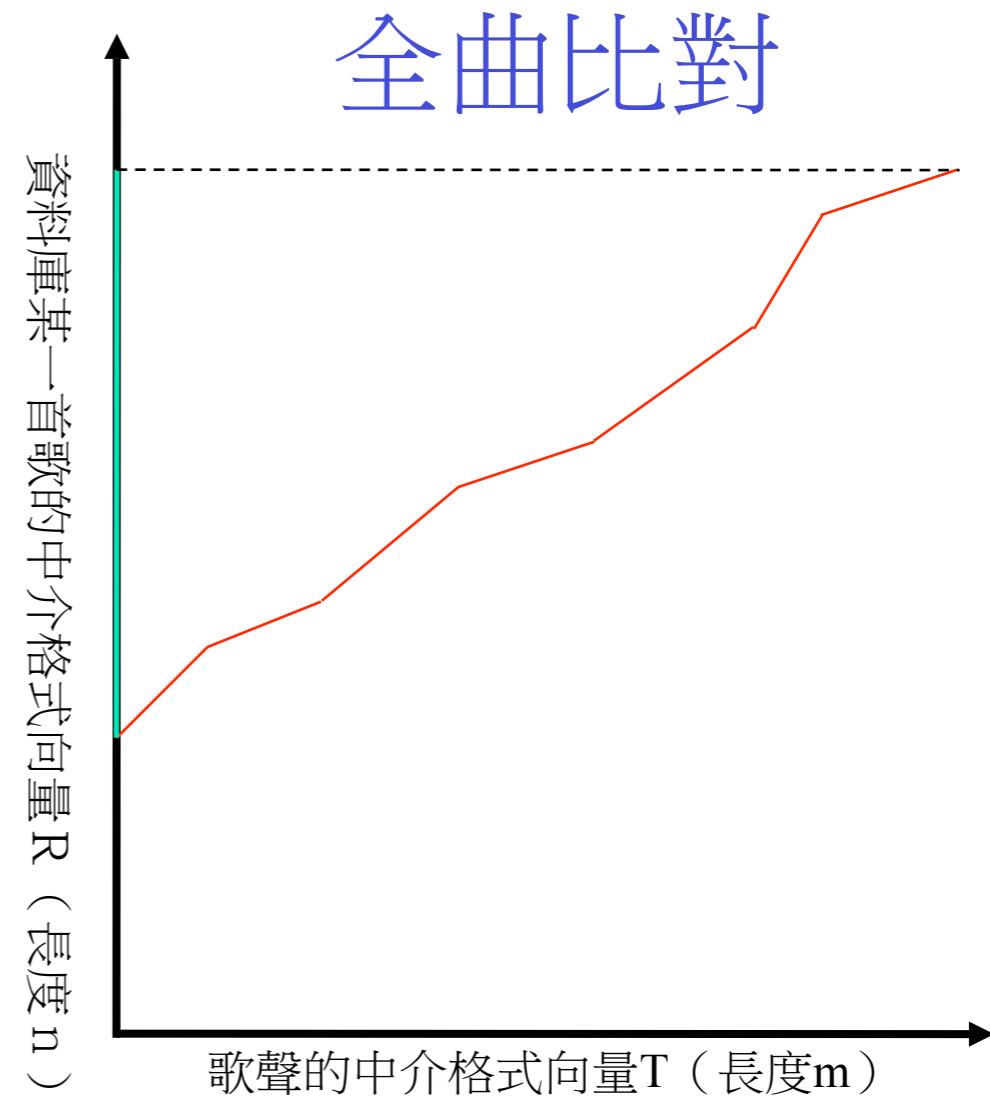
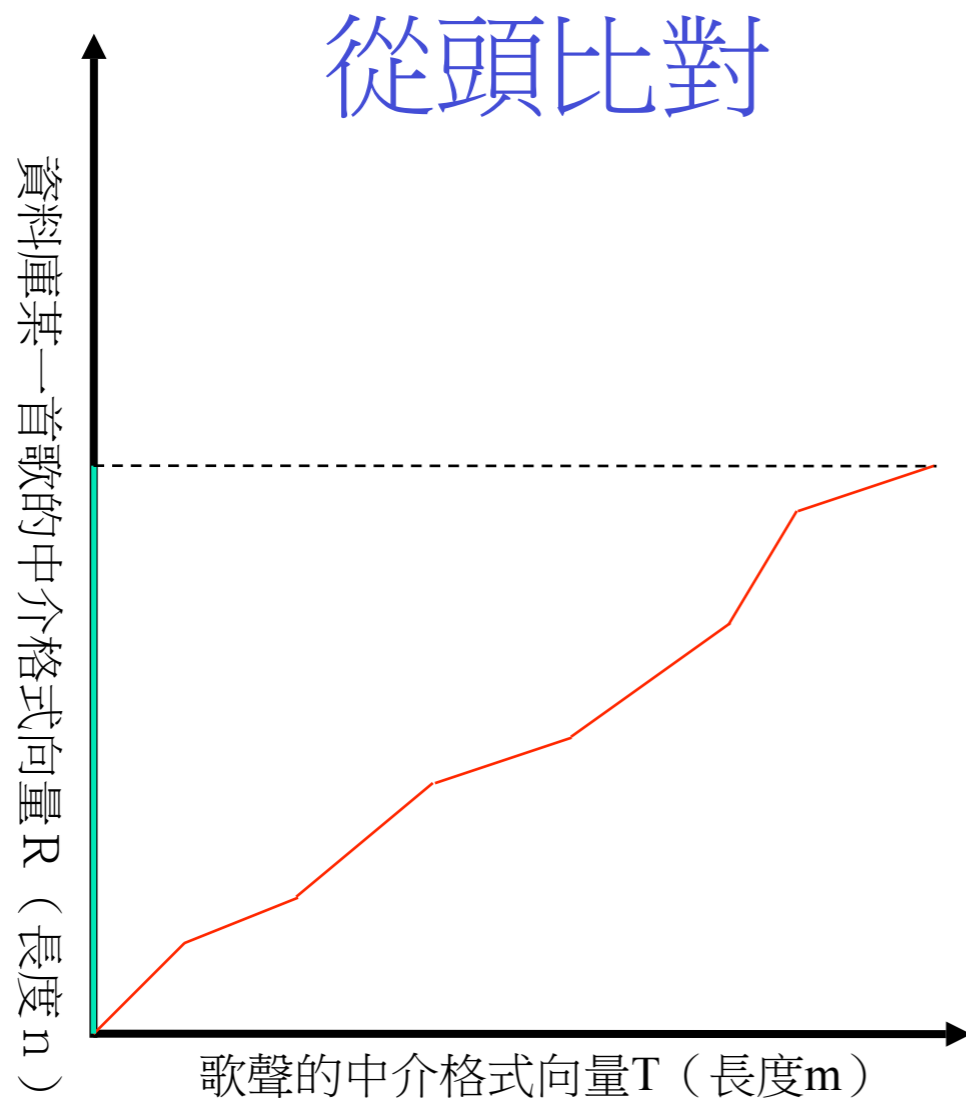
$$d(i, j) = \min \begin{cases} d(i-1, j-1) \\ d(i-2, j-1) + |T(i) - R(j)| \\ d(i-1, j-2) \end{cases}$$

$$dtw(T, R) = \min_{1 \leq j \leq n} d(m, j)$$

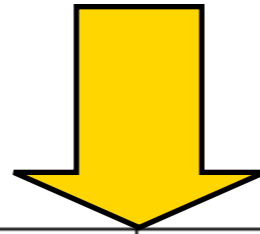
Dynamic Time Warping



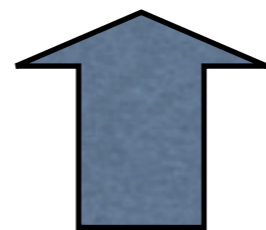
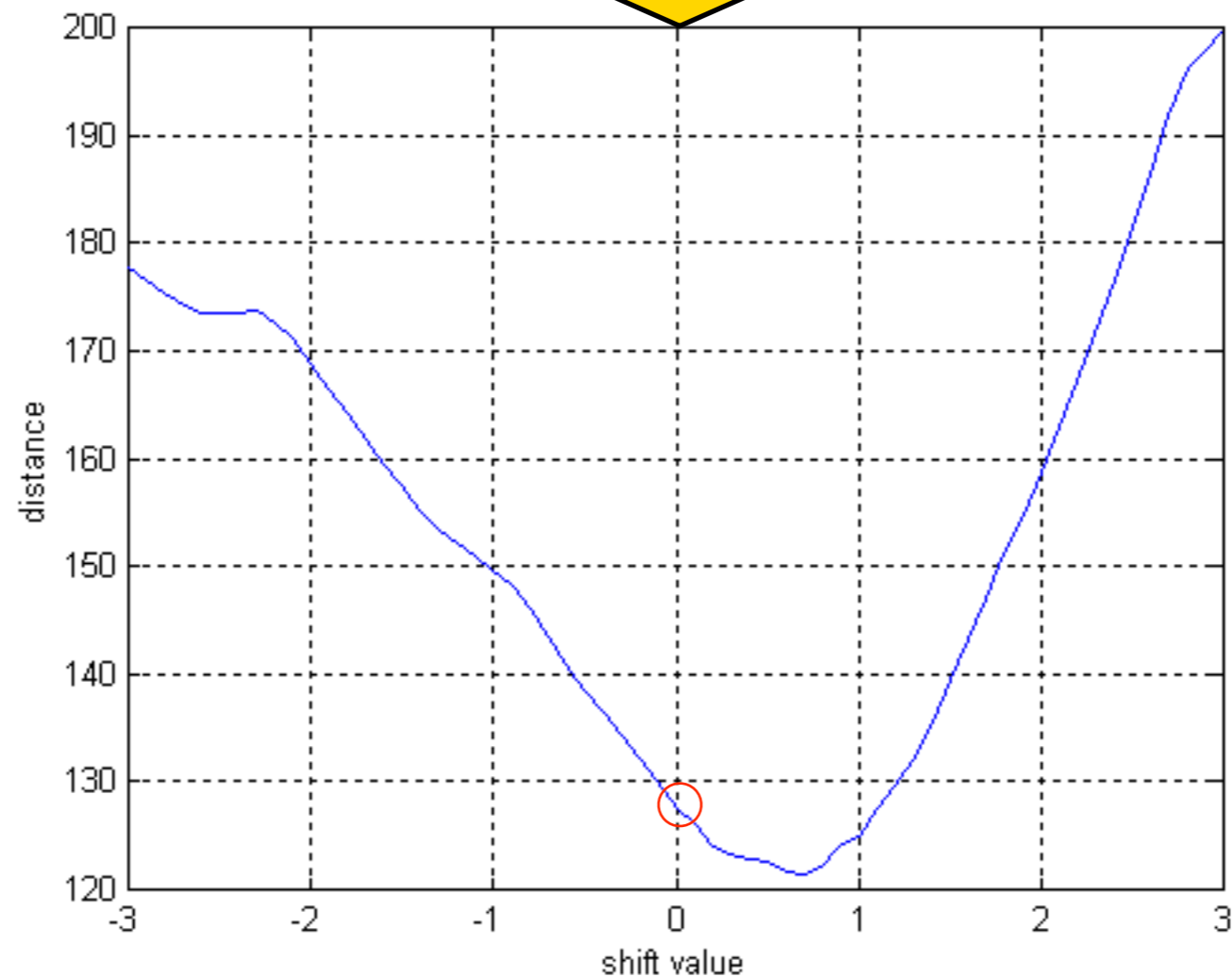
Dynamic Time Warping



Key Transposition



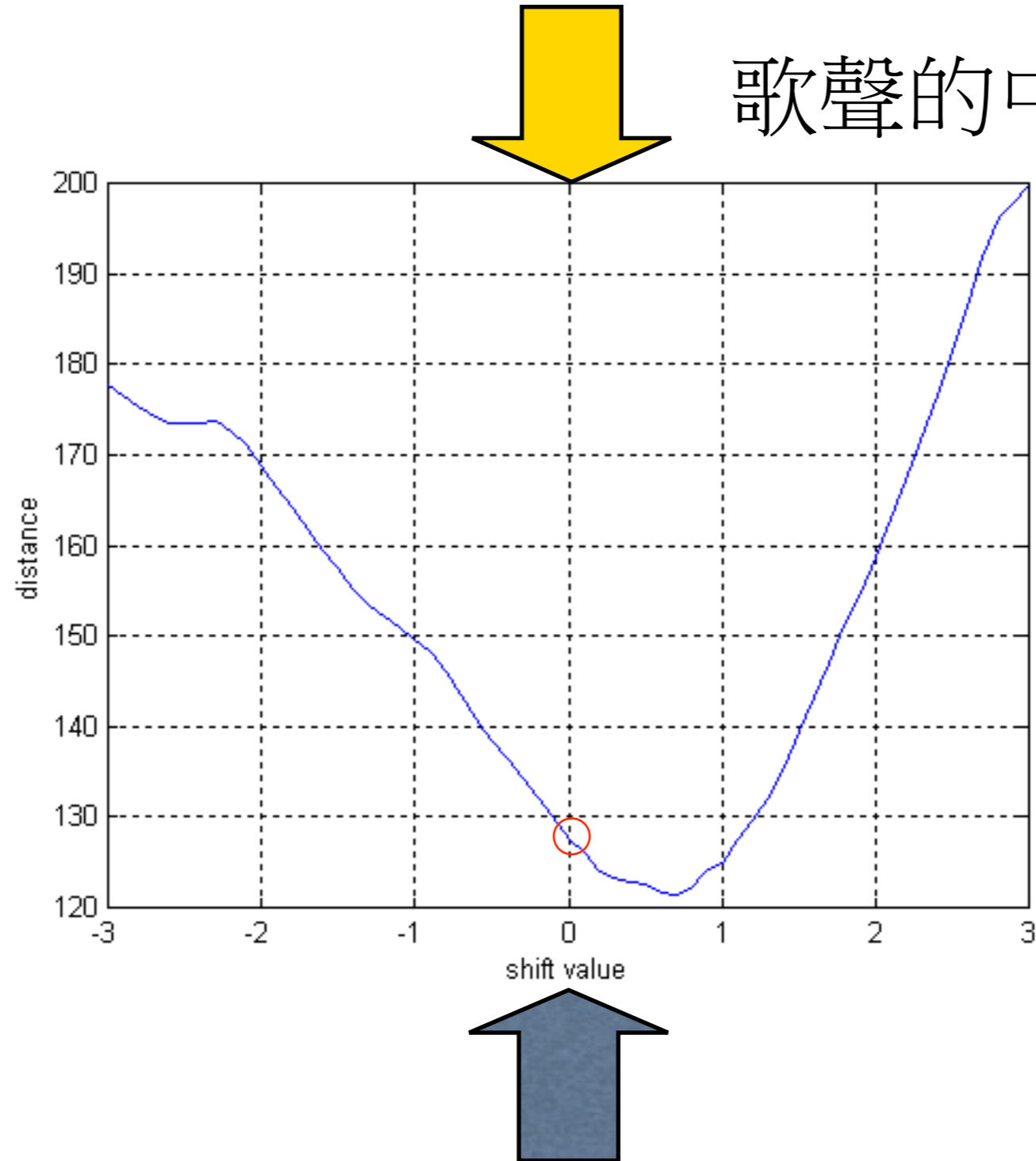
歌聲的中介格式



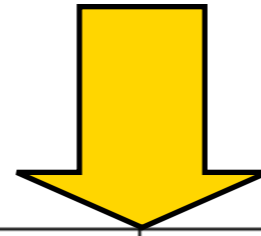
資料庫中某一首歌
的中介格式

Key Transposition

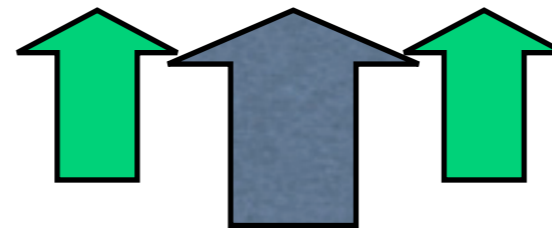
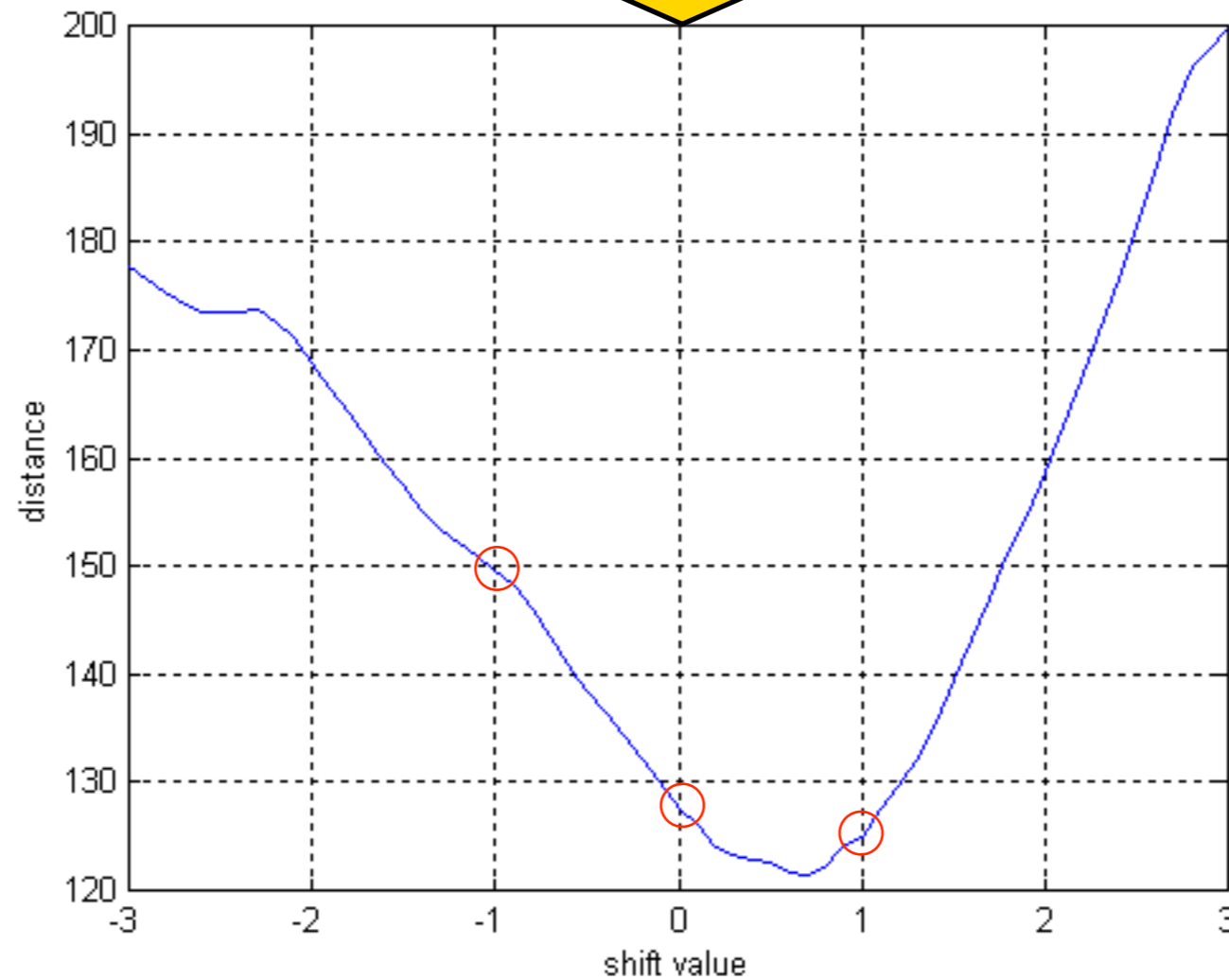
歌聲的中介格式



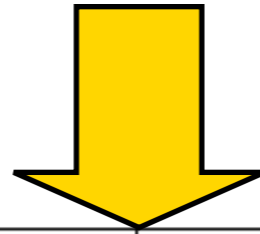
Key Transposition



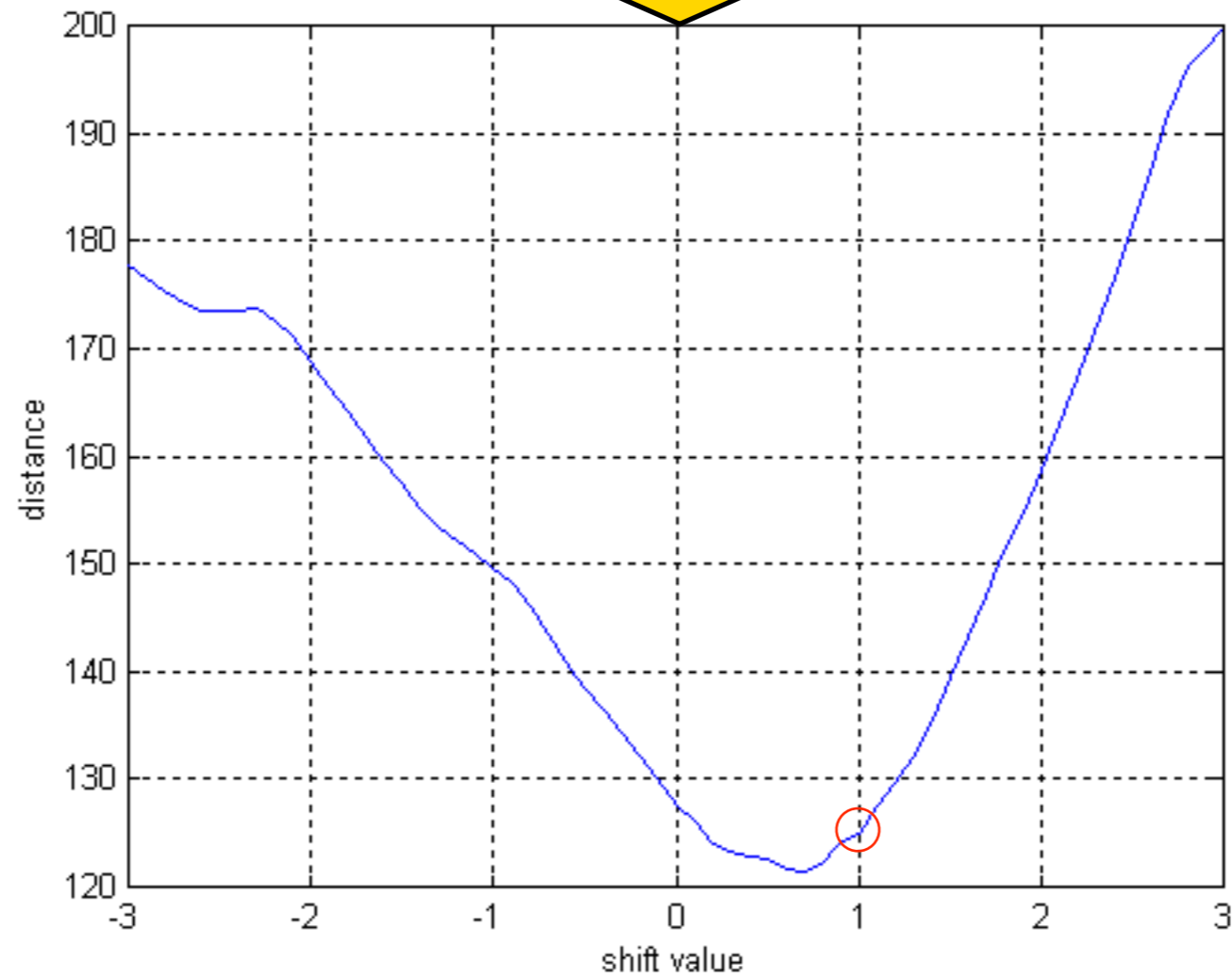
歌聲的中介格式



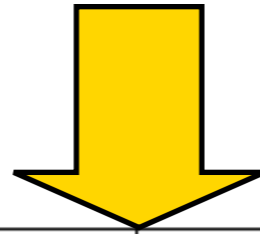
Key Transposition



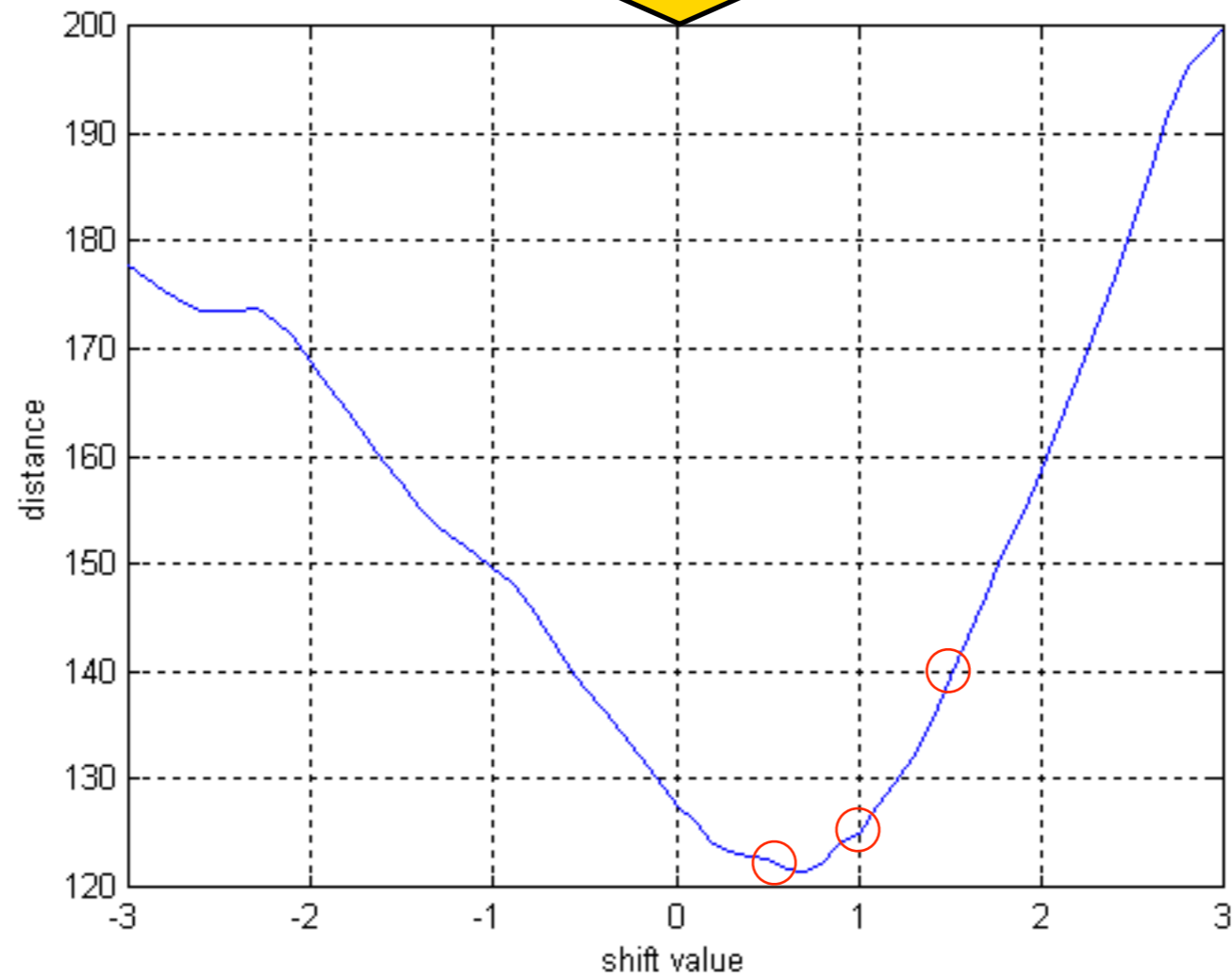
歌聲的中介格式



Key Transposition

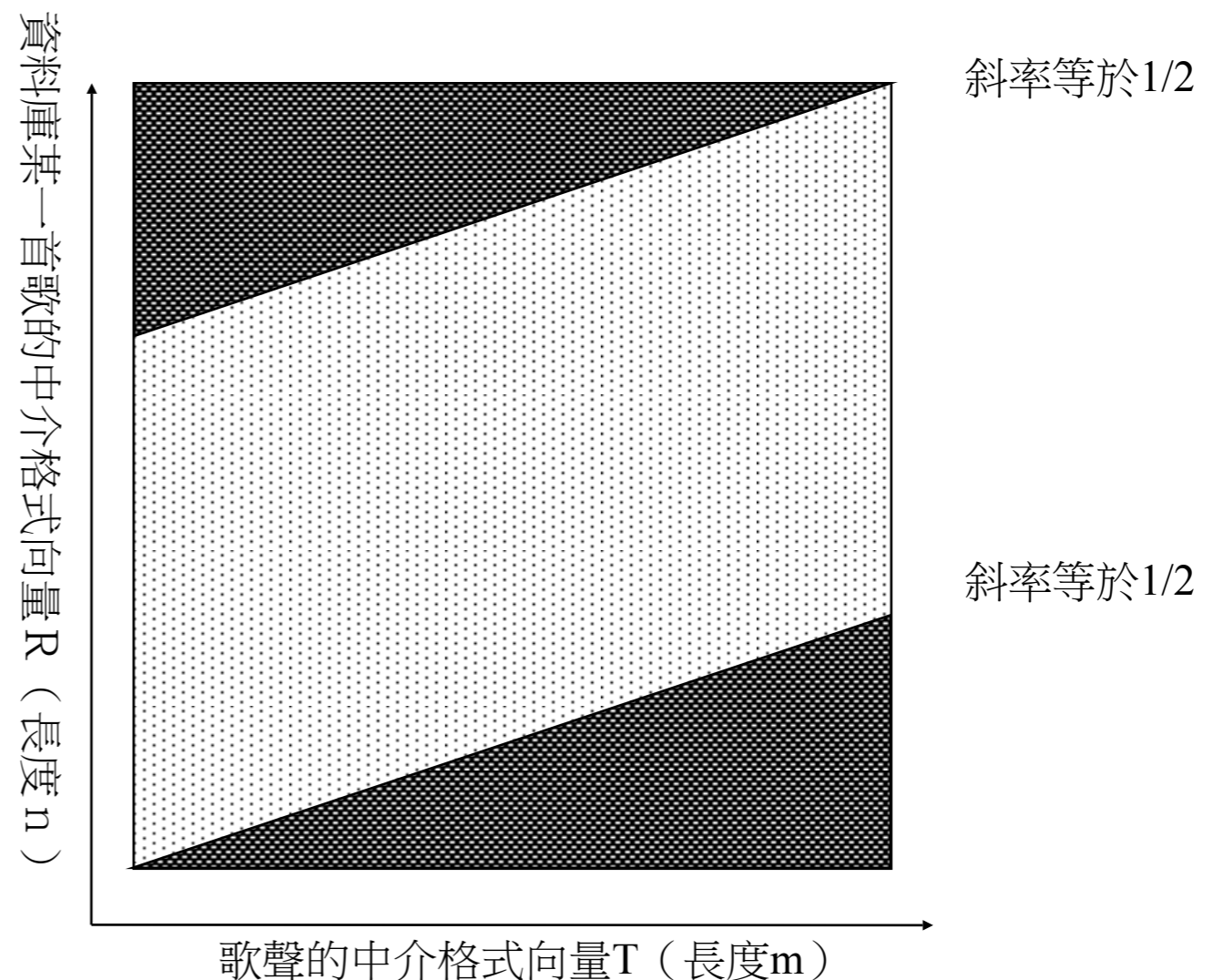


歌聲的中介格式



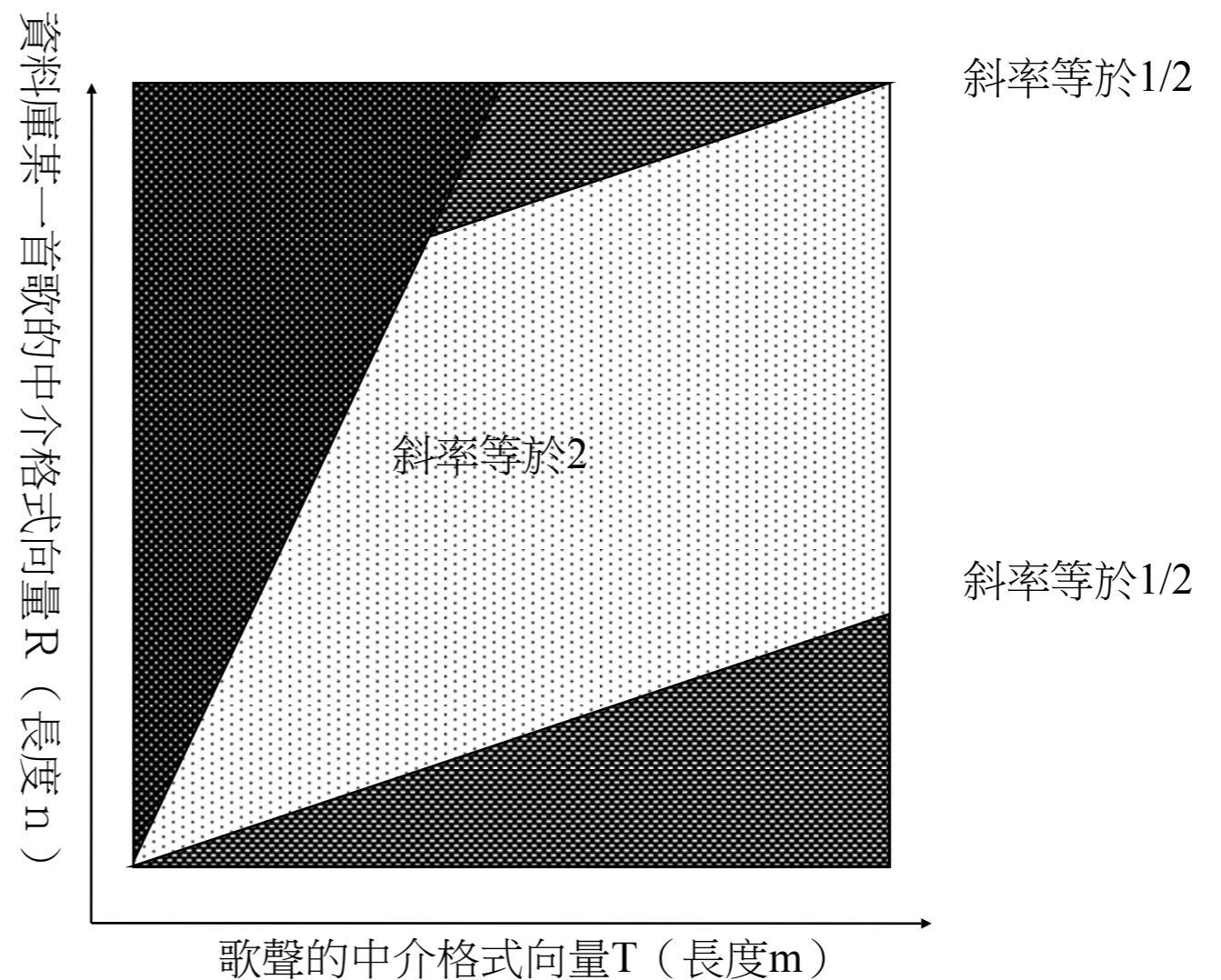
改進方法一

- 改用整數運算
 - 犧牲部分准确率
- 改良式dtw
 - 因為T的頭尾必須match到R的某一段，所以顏色較深的地方不必計算



改進方法一

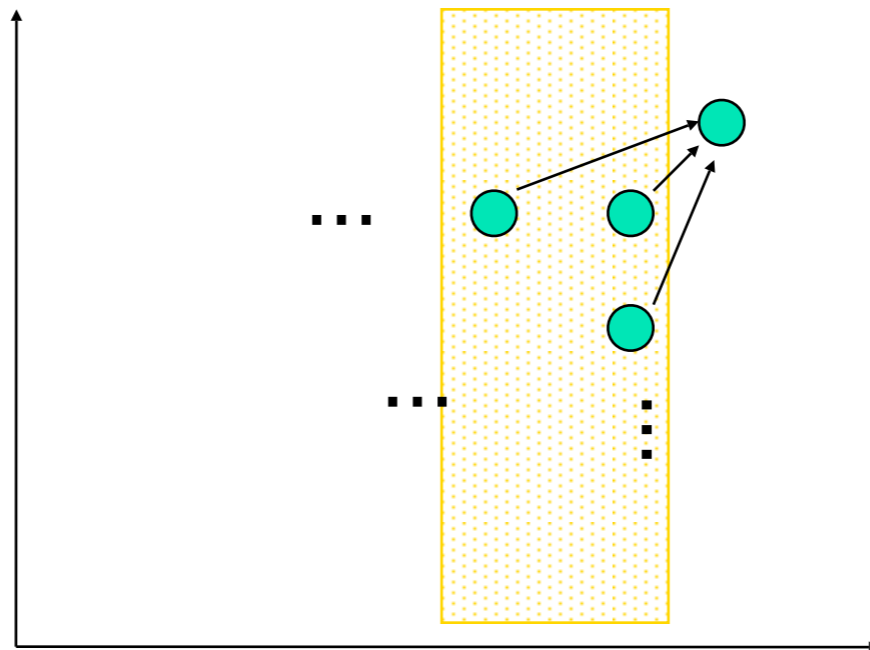
- 改用整數運算
 - 犧牲部分准确率
- 改良式dtw
 - 因為T的頭尾必須match到R的某一段，所以顏色較深的地方不必計算



改進方法一

- 改良式dtw

- 用dtw計算每一首歌的距離時，若發現dtw table最近兩行(column)的最小值超過之前最短距離的前k名，則停止dtw





改進方法二

- 將資料庫中每一首歌的中介格式，從每一個音符為起點切成數個長度為 **D=72** 的片段

中介格式 (69, 69, 67, 67, 67, 71, 72,)

→ 片段1 (69, 69, 67, 67, 67, 71, 72, ...)

片段2 (67, 67, 67, 71, 72, ...)

片段3 (71, 72, ...)

.....



改進方法二

- 用兩階段的方法比對

- 第一階段：線性伸縮比對（**linear scaling**）

- 將歌聲的中介格式伸縮**11**次（長度為原來的**0.75**倍到**1.25**倍），分別取出前 **D** 點後為 T_i （ $1 \leq i \leq 11$ ），假設資料庫中的第 **j** 首歌的中介格式有 n_j 個片段為 R_{jk} （ $1 \leq j \leq$ 資料庫中的歌曲數目， $1 \leq k \leq n_j$ ），令 T_i 和 R_{jk} 的距離為

$$\text{dist}(T_i, R_{jk}) = \sqrt{\sum_{t=1}^D \left((T_{it} - \bar{T}_i) - (R_{jkt} - \bar{R}_{jk}) \right)^2}$$



改進方法二

- 令資料庫第 j 首歌的分數為 $100 - \min_{\substack{1 \leq i \leq 11 \\ 1 \leq k \leq n_j}} \frac{dist(T_i, R_{jk})}{D}$ ，令

$$k' = \arg \min_{\substack{1 \leq i \leq 11 \\ 1 \leq k \leq n_j}} dist(T_i, R_{jk})$$

- 篩選出前 $n=200$ 首分數最高者做第二階段的比對
- 缺點：每和資料庫中第 j 首歌做比對就要計算 $n_j * 11$ 次距離
- 第二階段：dynamic time warping
 - 將篩選後的每一首歌的最接近片段平移音調4次，總共和歌聲原本的中介格式計算5次距離，找出最小值並轉換成分數，當成資料庫該首歌最後的分數



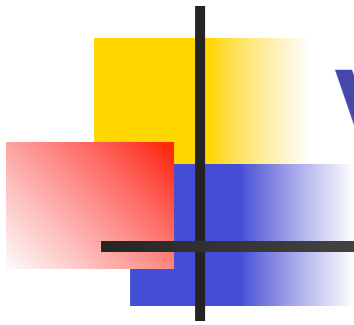
針對全曲比對的加速方法

- 用兩階段式比對
- 在第一階段將資料庫歌曲的每一個片段看成一個 D 度空間中的資料點，歌聲伸縮後的中介格式則是空間中的查詢點，利用快速找最近鄰居的方法找出每首歌最接近的片段
 - Vantage-point tree
 - Branch-and-bound tree
 - Equal-average hyperplane partitioning method

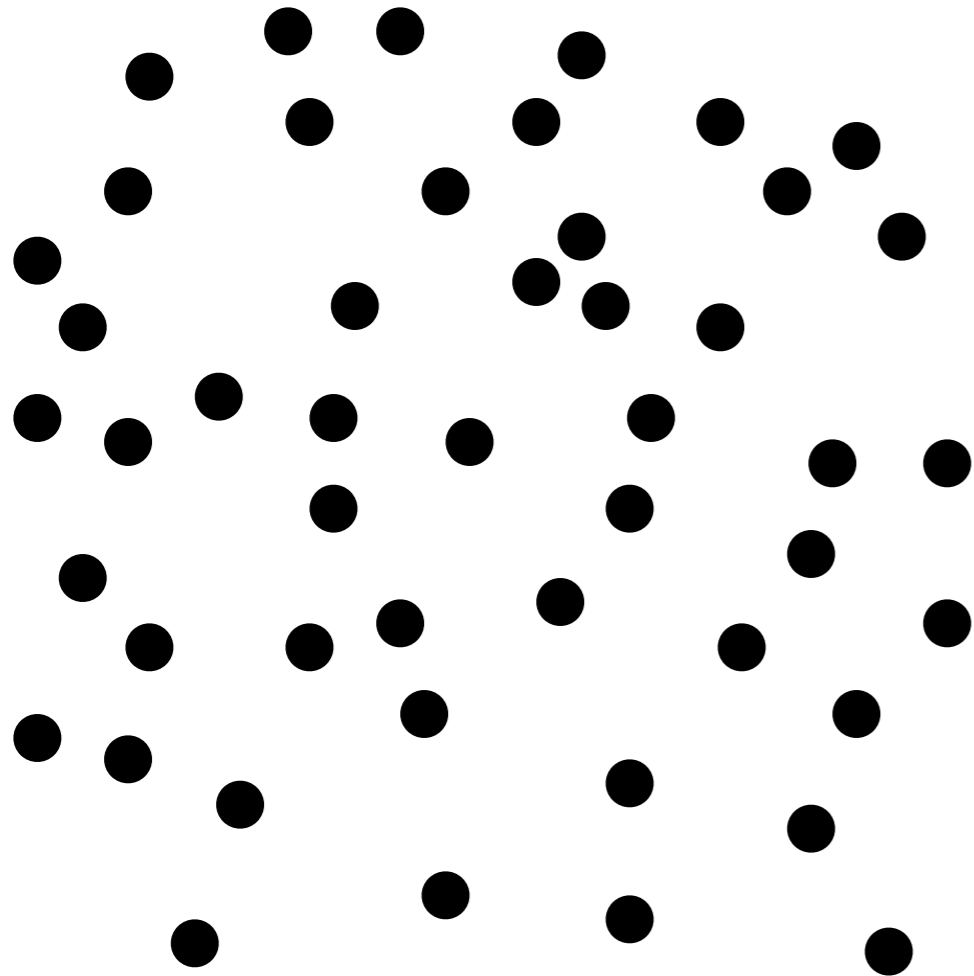


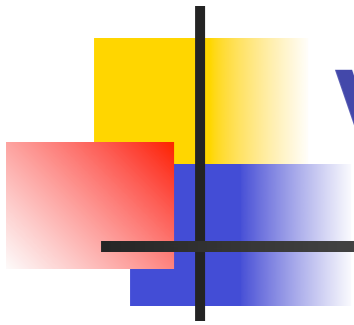
Vantage-Point Tree

- 將資料點建立成一個平衡的二元樹，利用距離滿足三角不等式的關係，可以減少計算查詢點到資料點的距離的次數
 - 缺點：建樹麻煩，使用時耗記憶體
 - 困難：資料點少，查詢點到所有資料點的距離都差不多，加速效果不顯著

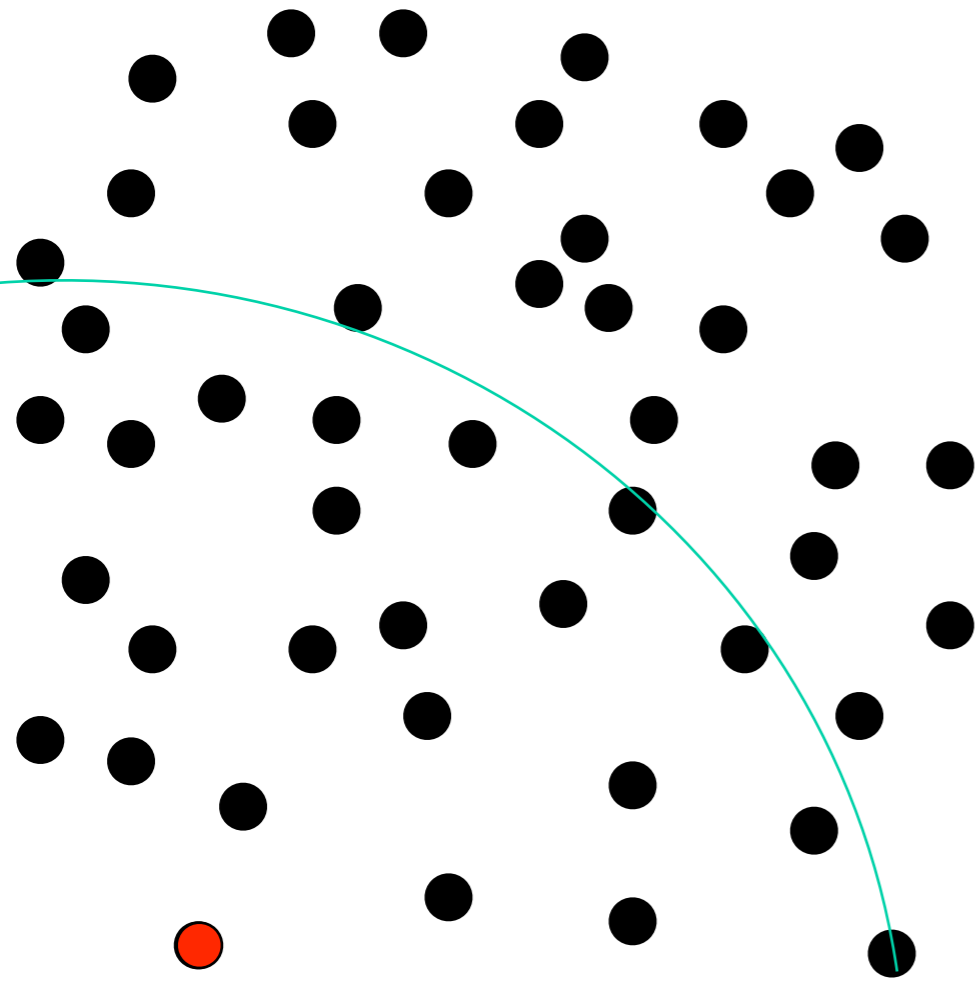


Vantage-Point Tree

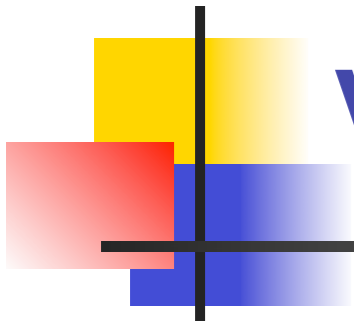




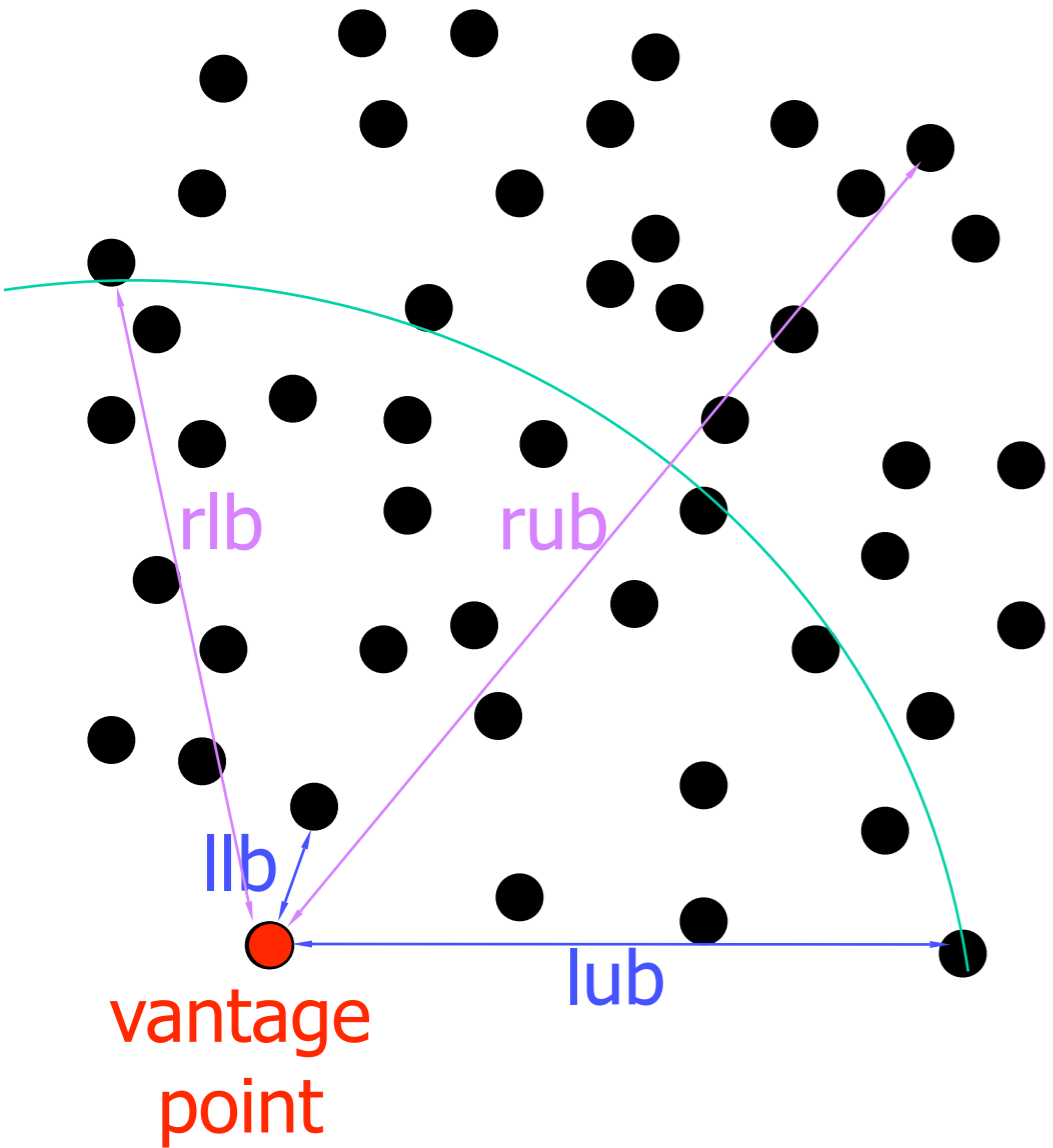
Vantage-Point Tree



vantage
point

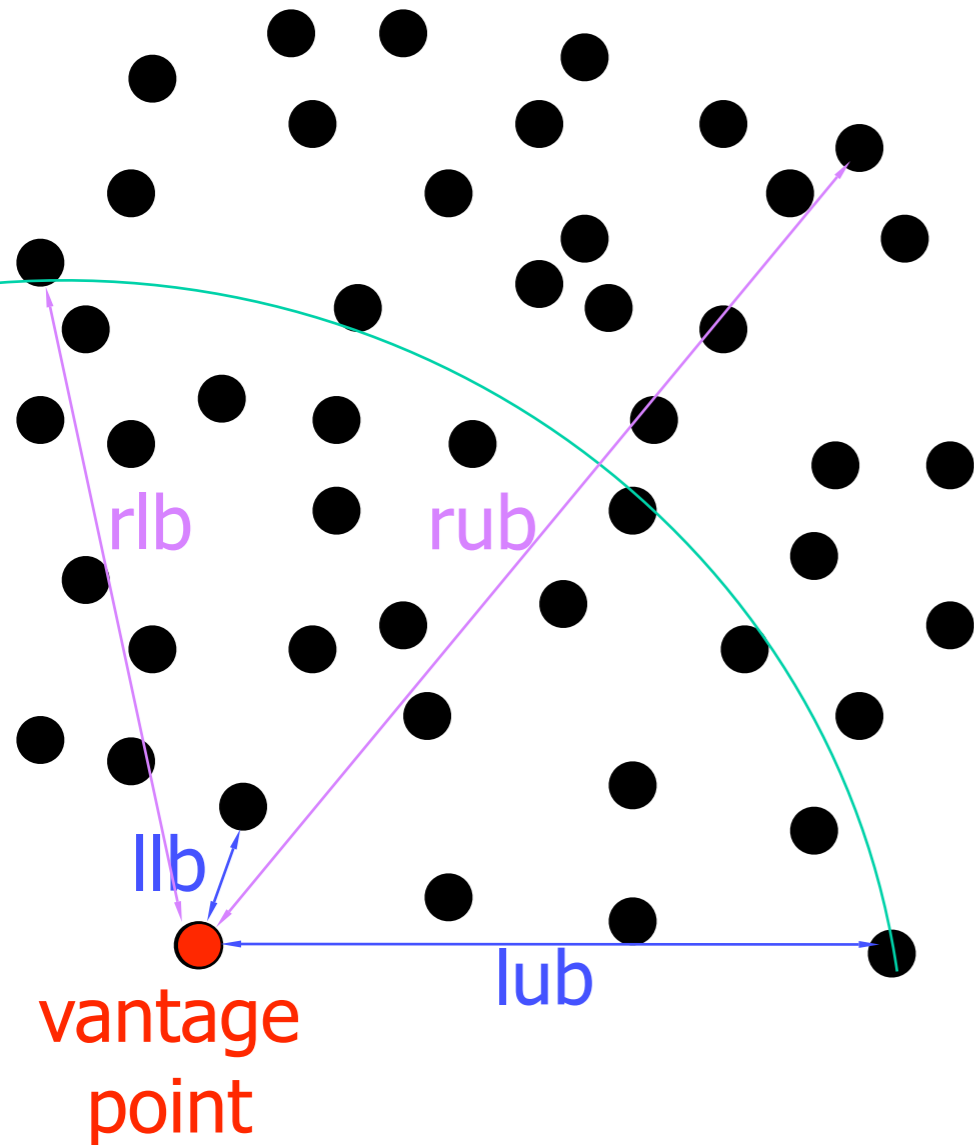
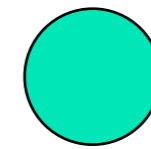


Vantage-Point Tree



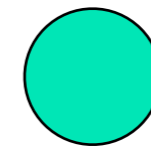
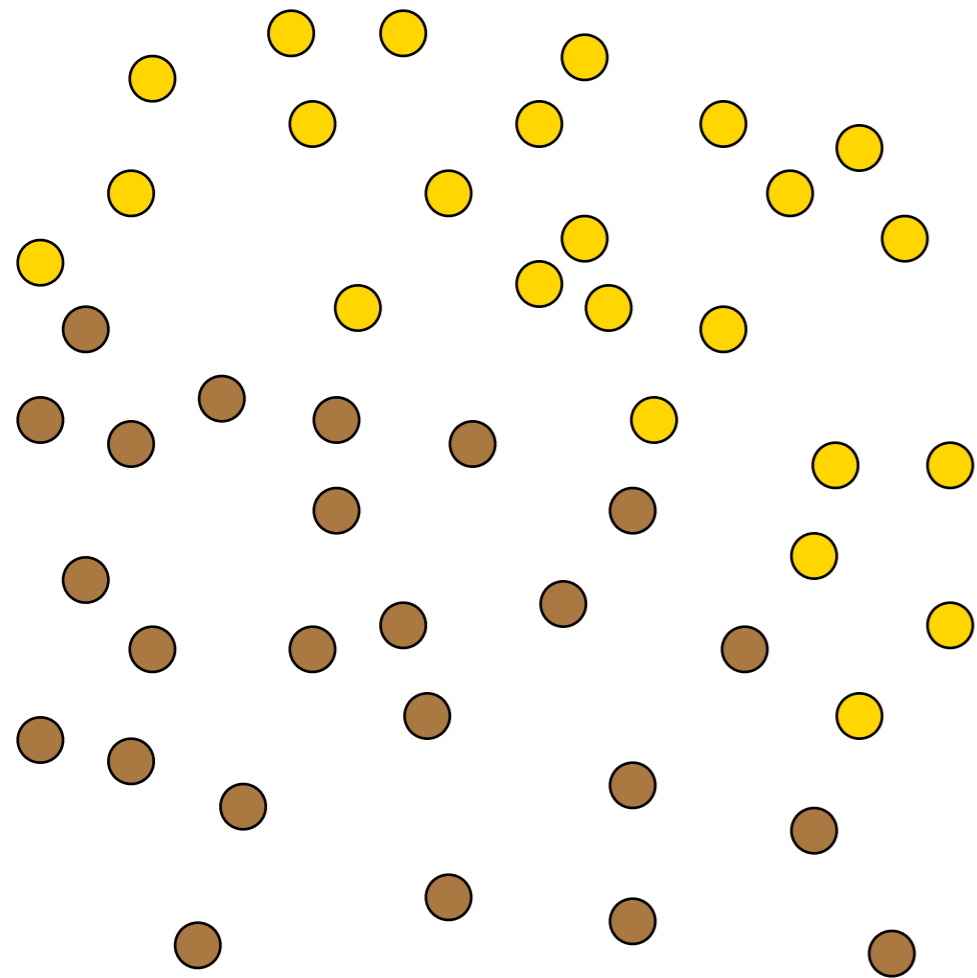
Vantage-Point Tree

Vantage point
Left lower bound
Left upper bound
Right lower bound
Right upper bound



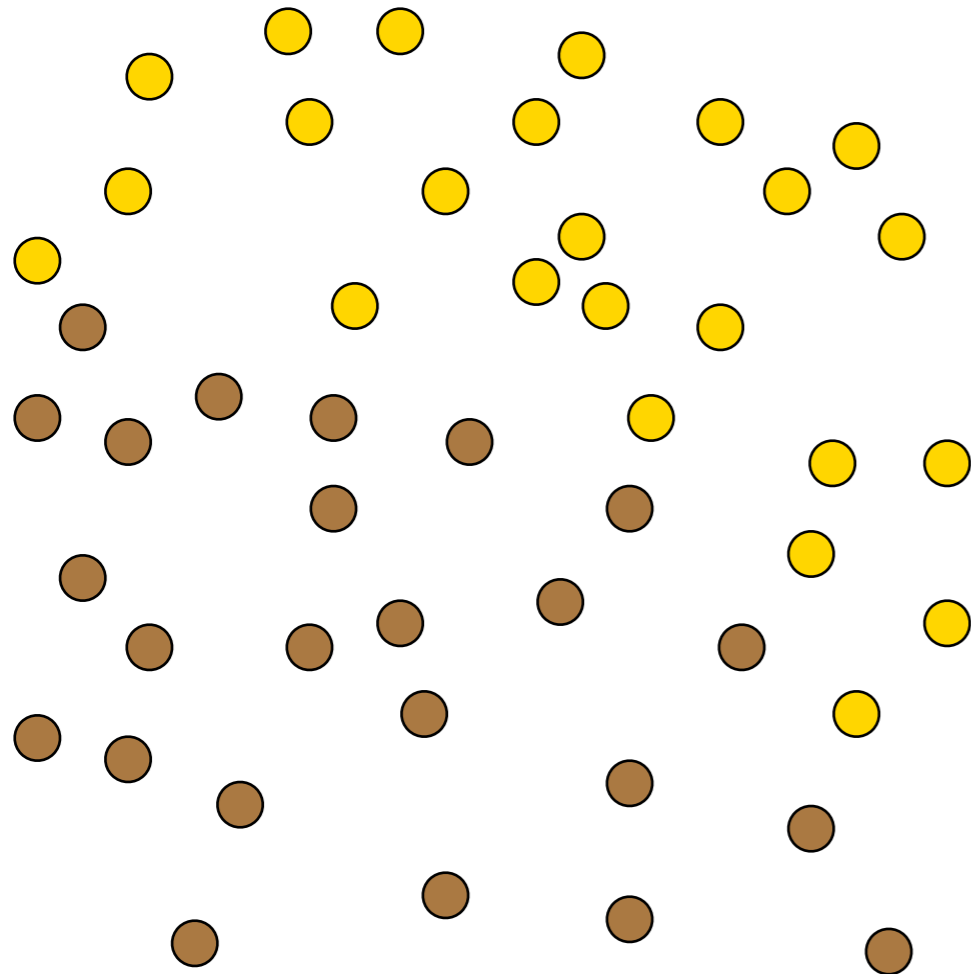
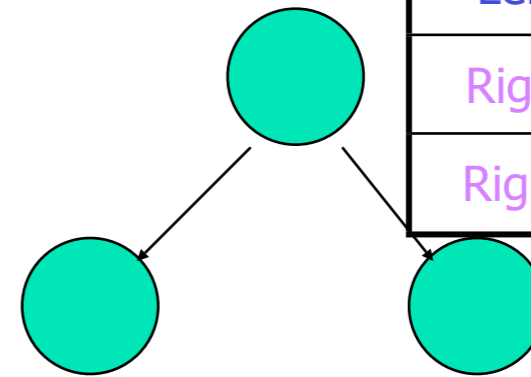
Vantage-Point Tree

Vantage point
Left lower bound
Left upper bound
Right lower bound
Right upper bound



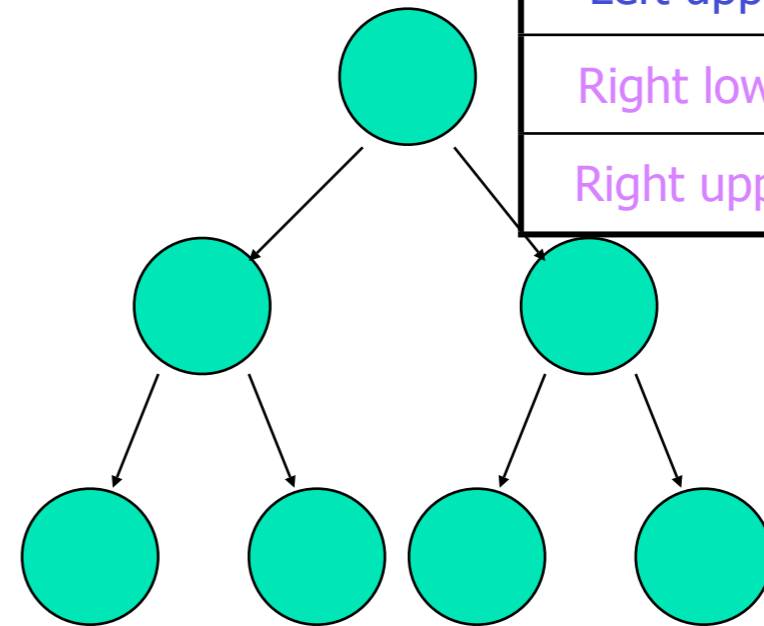
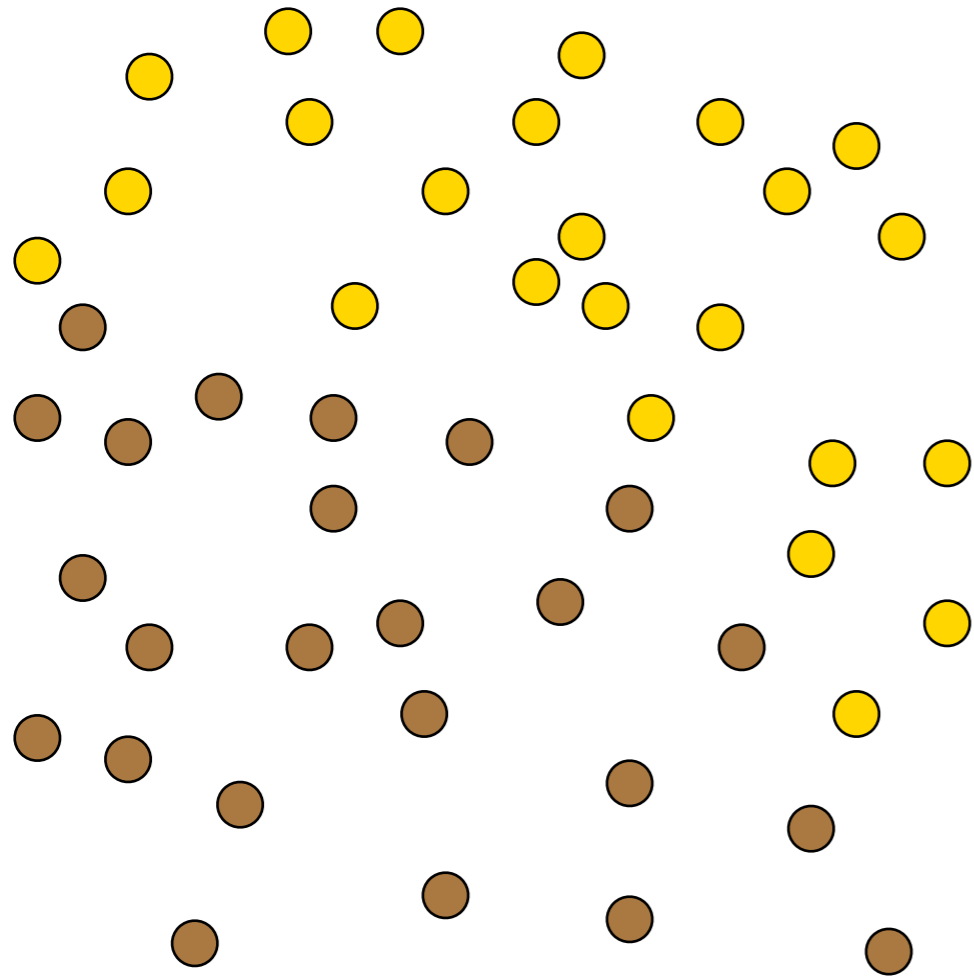
Vantage-Point Tree

Vantage point
Left lower bound
Left upper bound
Right lower bound
Right upper bound



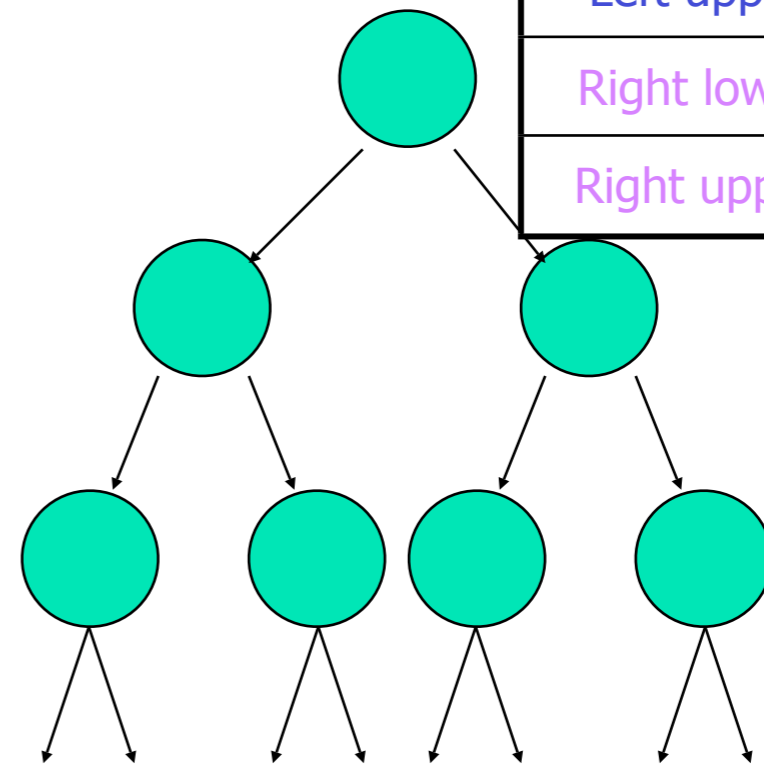
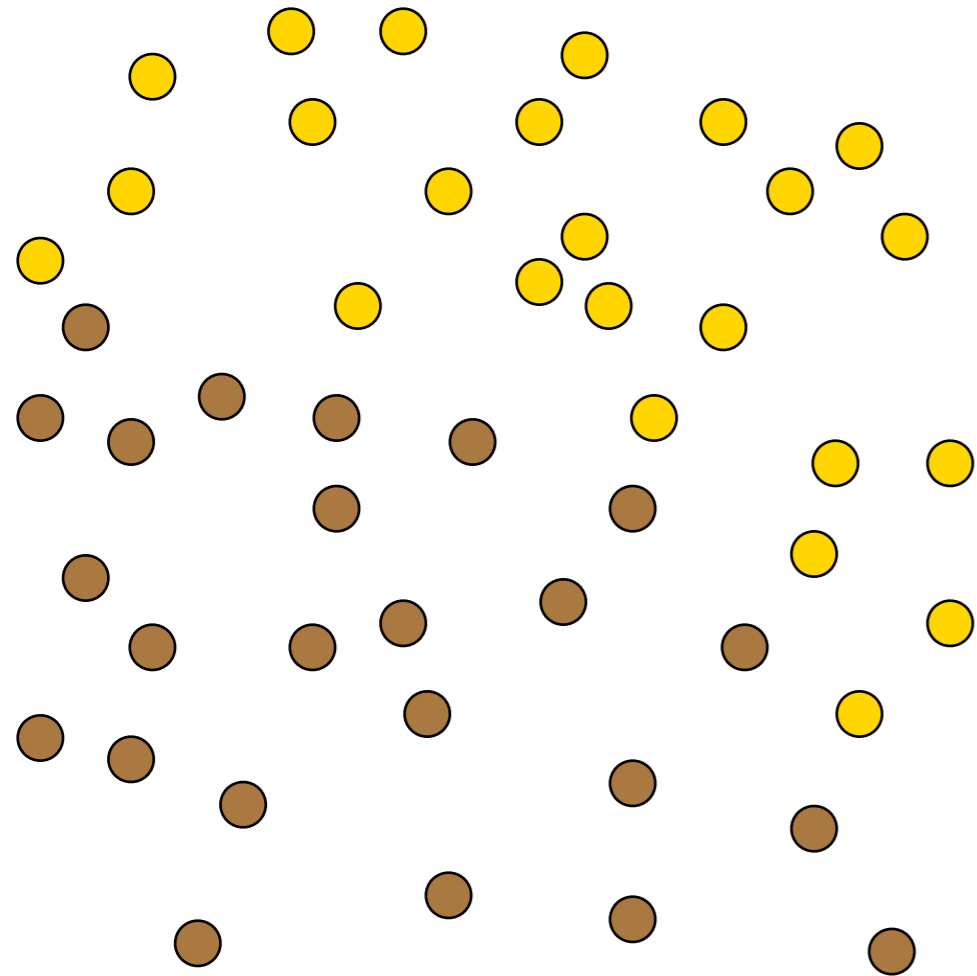
Vantage-Point Tree

Vantage point
Left lower bound
Left upper bound
Right lower bound
Right upper bound



Vantage-Point Tree

Vantage point
Left lower bound
Left upper bound
Right lower bound
Right upper bound



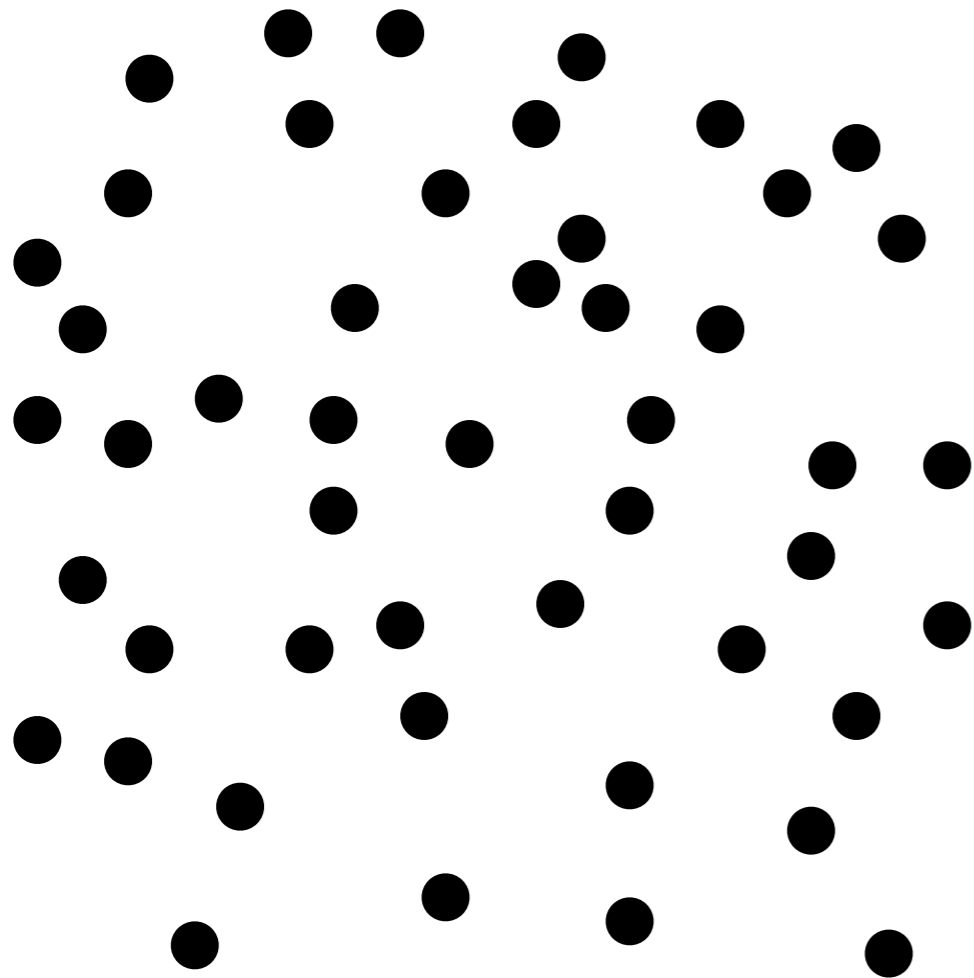


Branch-and-Bound Tree

- 原理類似 **vantage-point tree**，將資料點建立成一個多元樹，利用距離滿足三角不等式的關係，可以減少計算查詢點到資料點的距離的次數
 - 缺點：建樹時使用 **K-means**，麻煩耗時，使用時耗記憶體
 - 困難：資料點少，查詢點到所有資料點的距離都差不多，加速效果不顯著

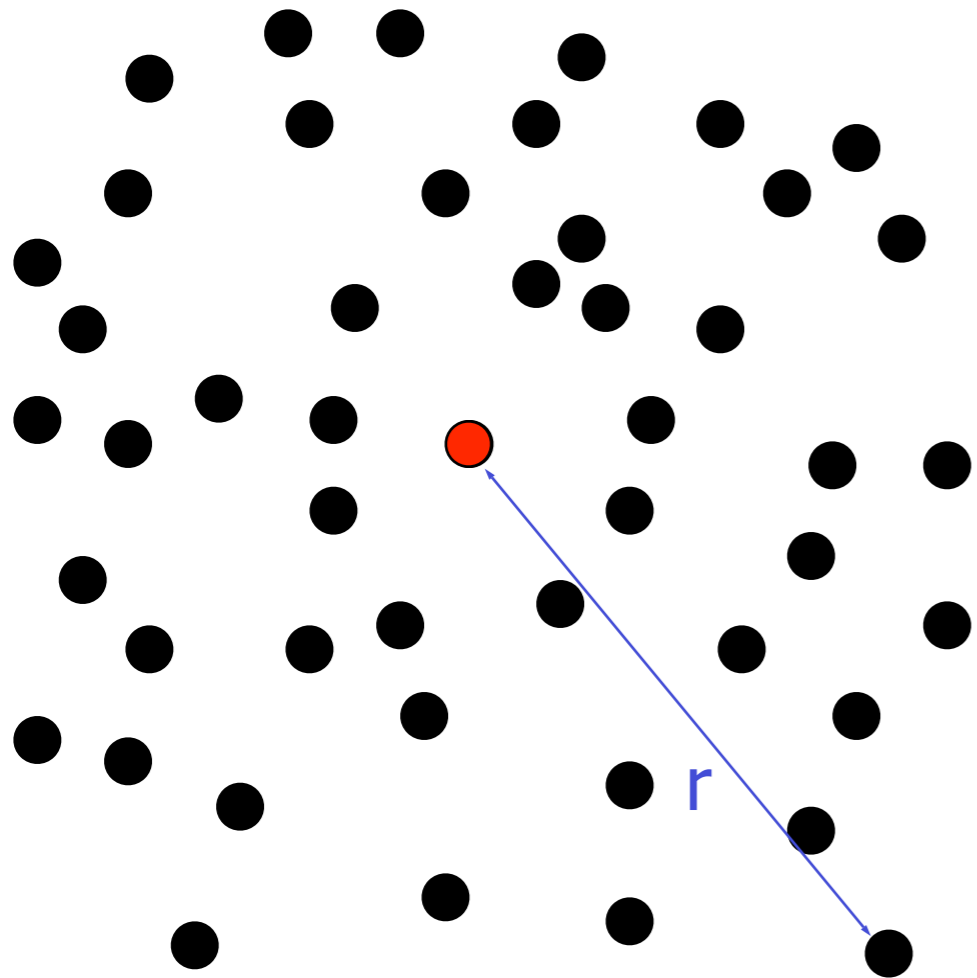


Branch-and-Bound Tree

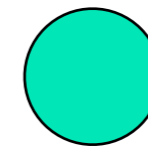
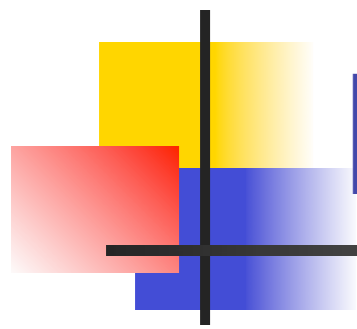




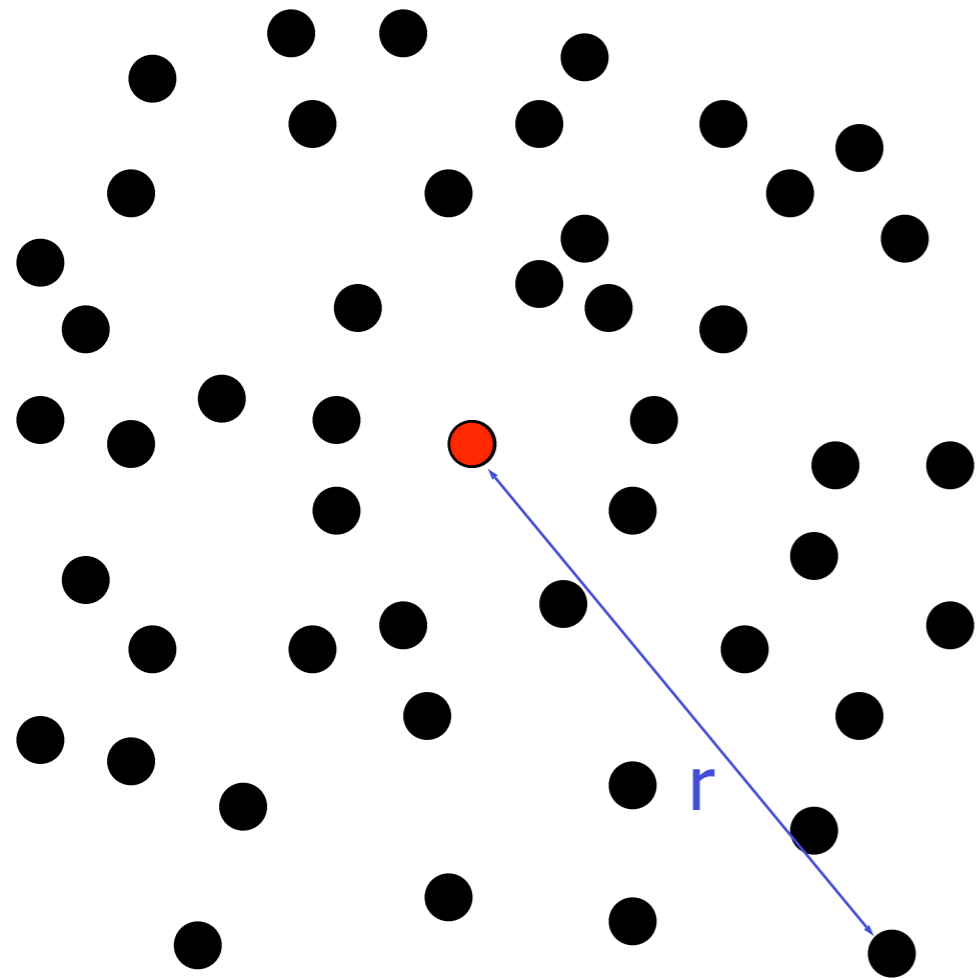
Branch-and-Bound Tree



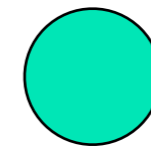
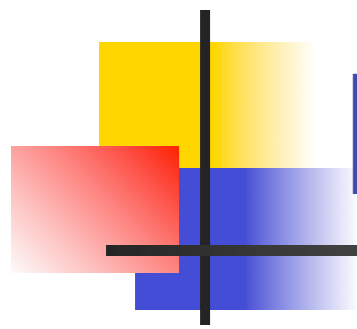
Branch-and-Bound Tree



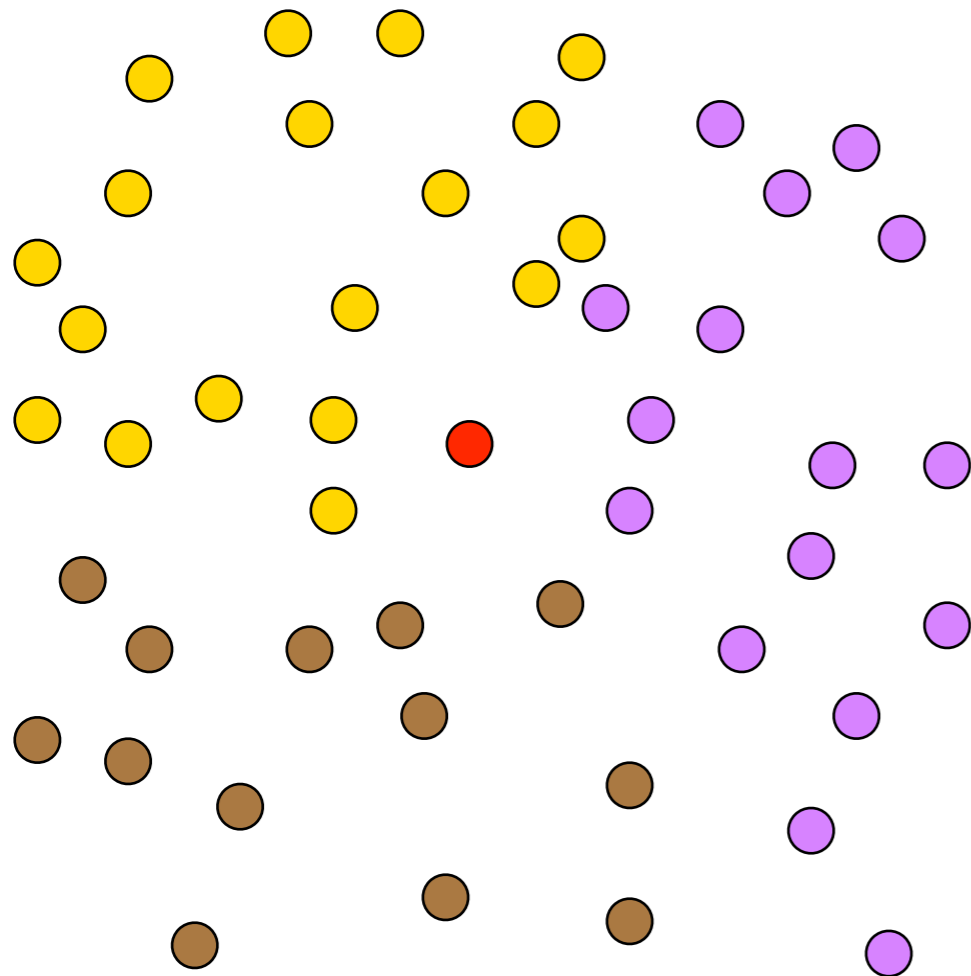
Center
Radius



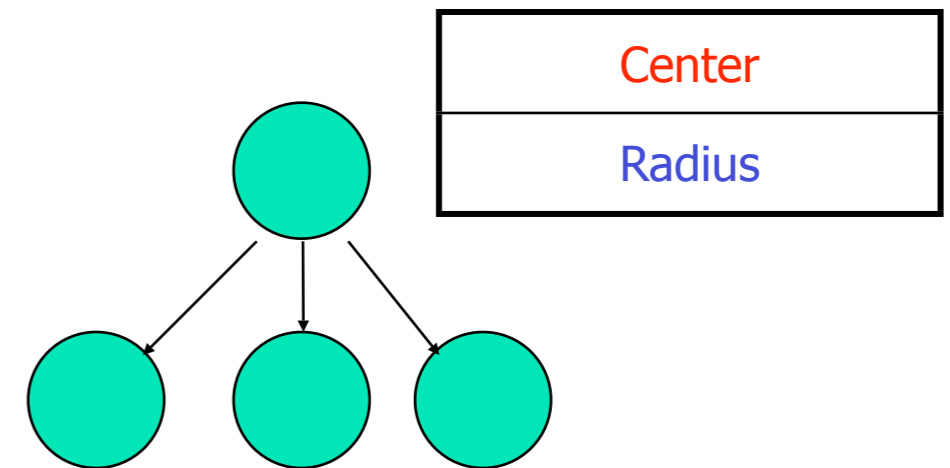
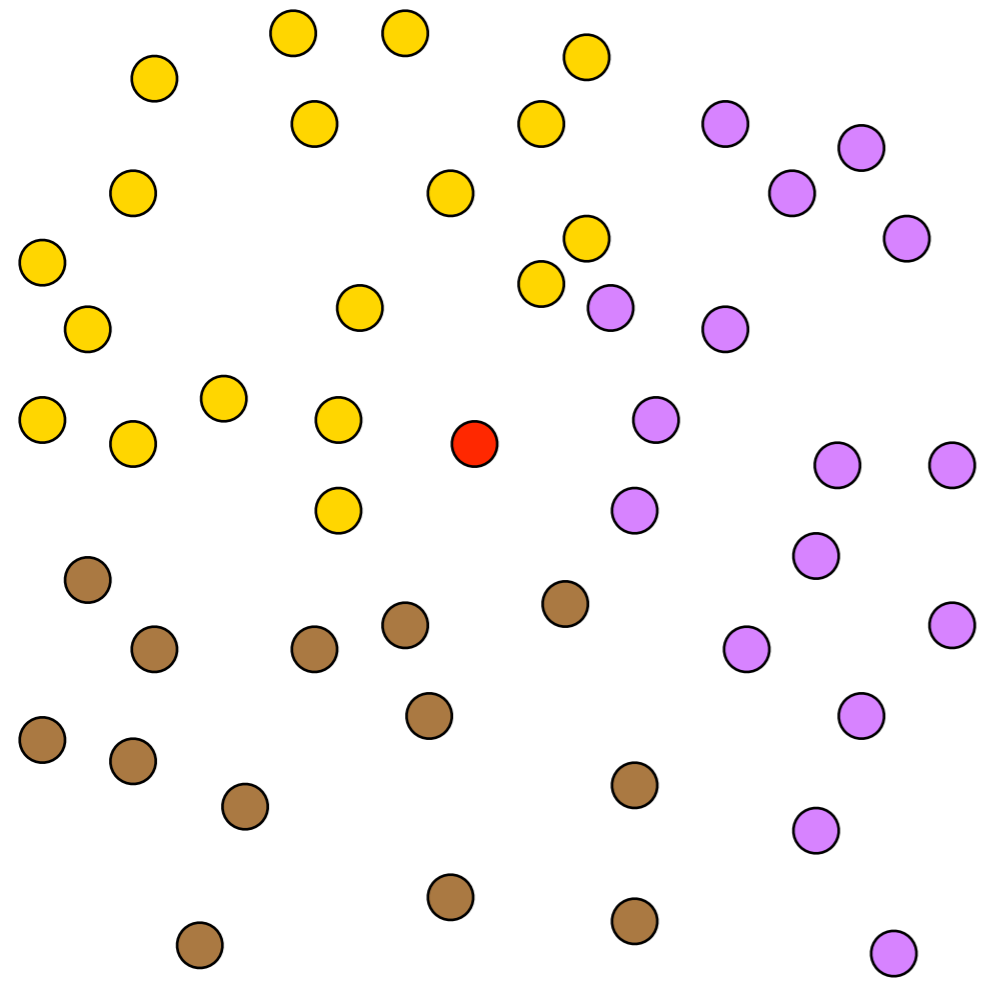
Branch-and-Bound Tree



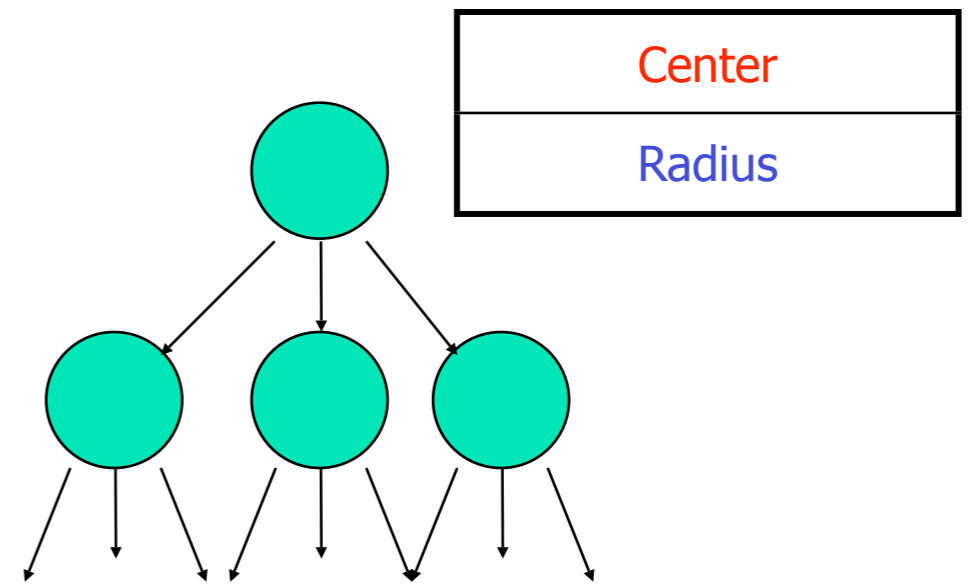
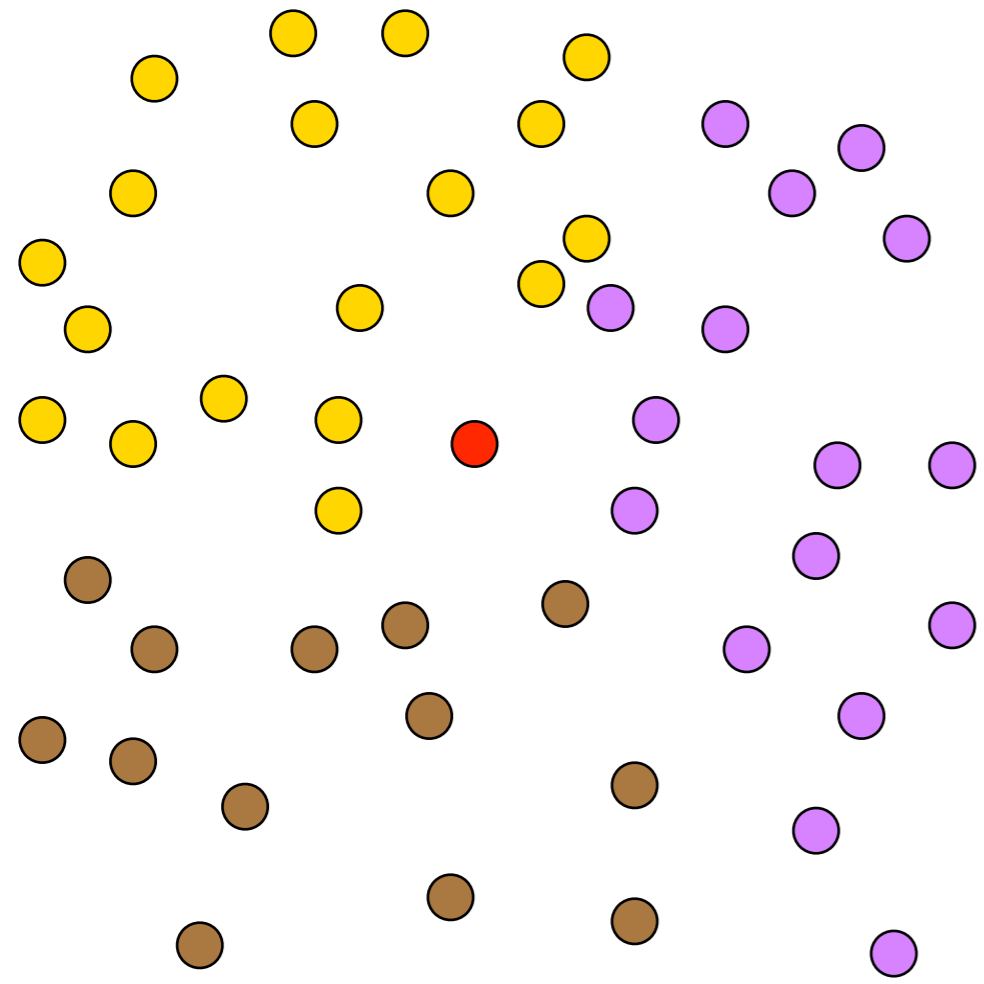
Center
Radius



Branch-and-Bound Tree



Branch-and-Bound Tree



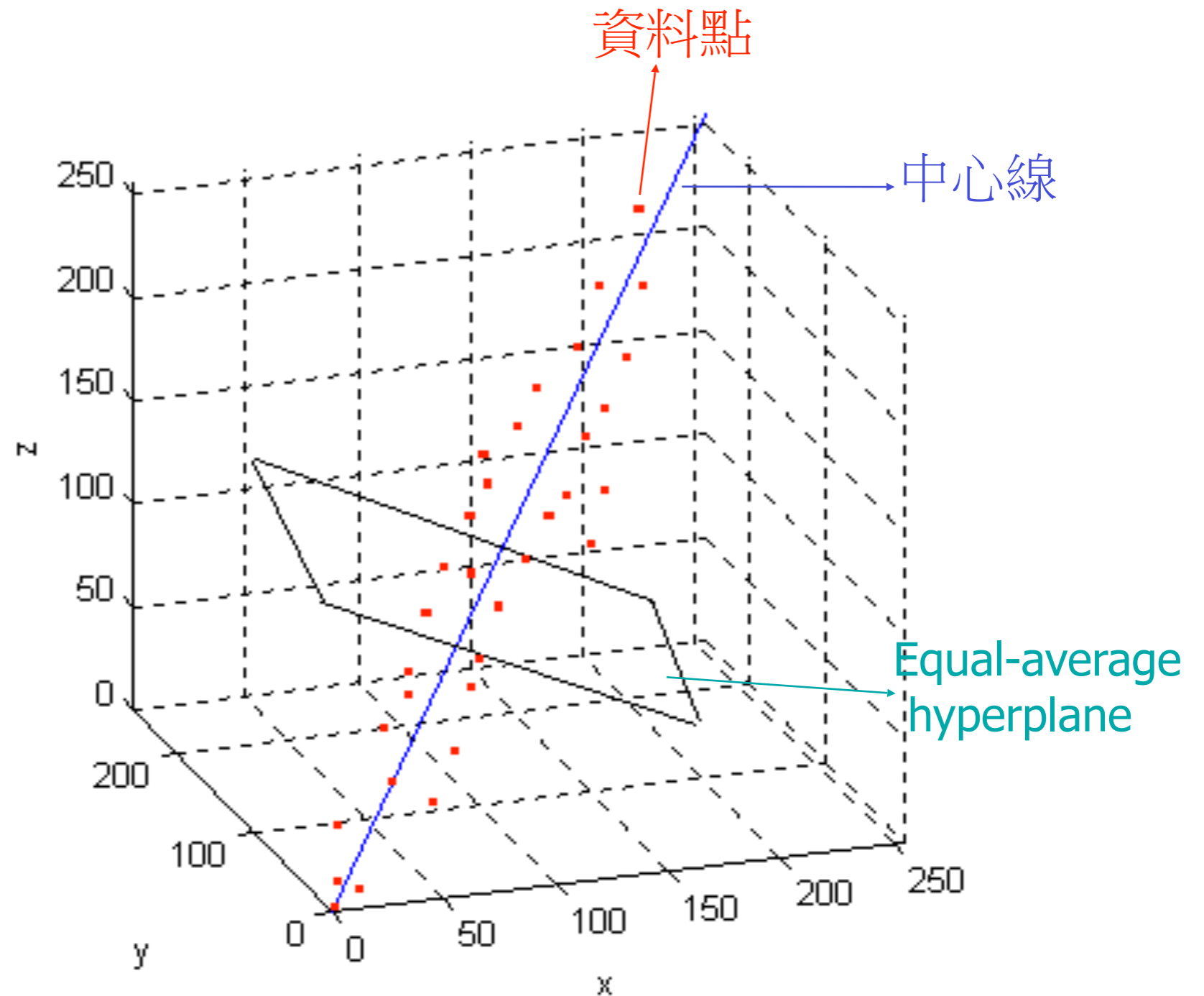


Equal-average Hyperplane Partitioning Method

- 讓資料點沿一直線分散開來，事先算出每一個資料點投影到直線上的位置。利用兩資料點間的距離會大於或等於這兩點投影到該直線上的距離的關係，可以減少計算查詢點到資料點的距離的次數
 - 優點：實作容易，省記憶體
 - 缺點：使用時機較狹隘

Equal-average Hyperplane Partitioning Method

- 中心線 L 是一條通過原點和 $(1, 1, 1, \dots, 1)$ 的直線
- Equal-average hyperplane是和 L 垂直的面



Equal-average Hyperplane Partitioning Method

- 假設 $\mathbf{a}=(s, s, \dots, s)$ 為 L 上的一個點
- 則和 L 相交於 \mathbf{a} 點的 equal-average hyperplane 其方程式為

$$H(x) = s \sum_{i=1}^D x_i - Ds^2 = 0$$

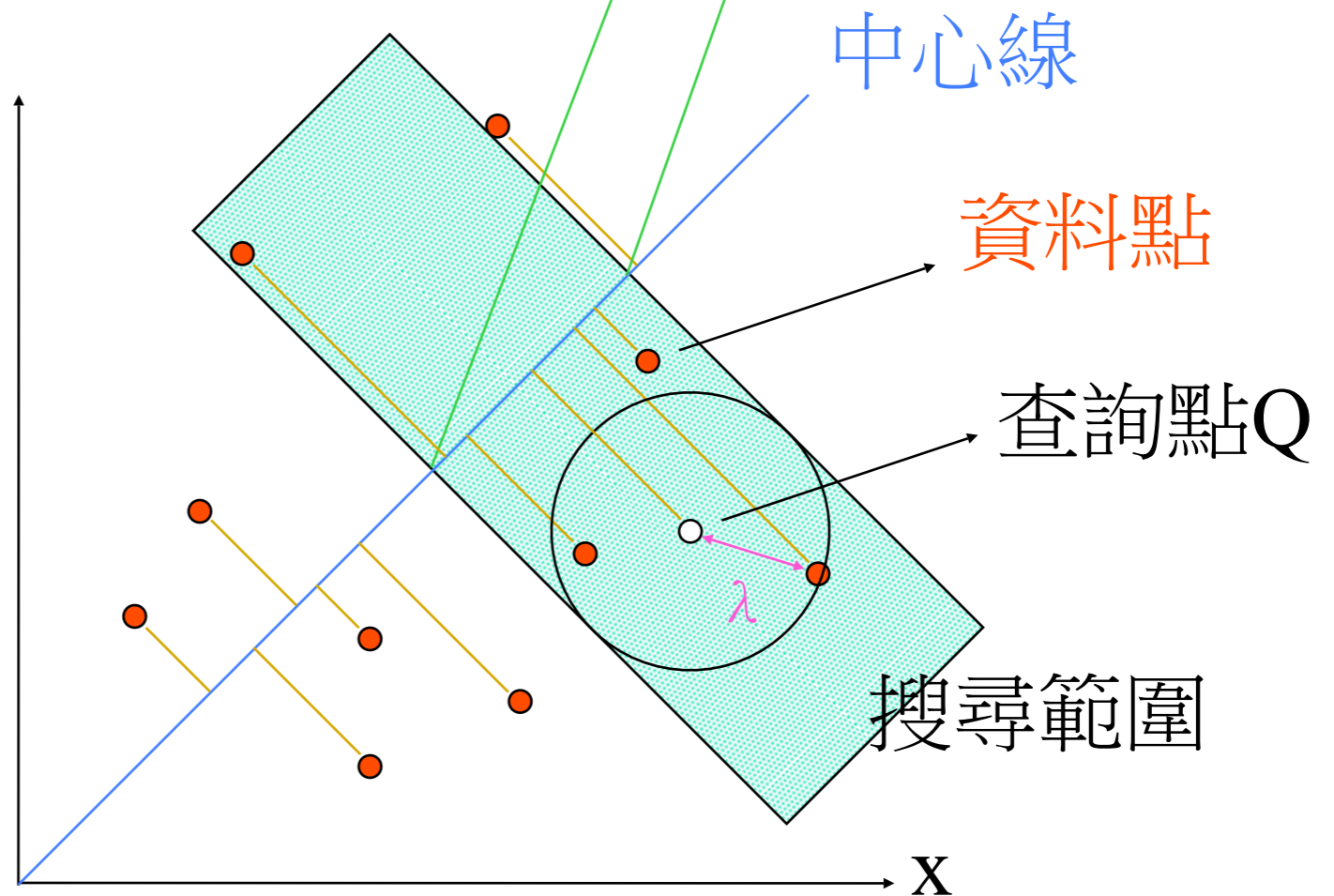
$$\Rightarrow s = \frac{1}{D} \sum_{i=1}^D x_i \quad \leftarrow \text{座標平均值}$$

- 座標平均值相同的點會落在同一個 equal-average hyperplane

Equal-average Hyperplane Partitioning Method

- 座標平均值落在 $[S_{\min}, S_{\max}]$ 之外的點不可能是最近鄰居

$$S_{\min} = \frac{1}{D} \left(\sum_{i=1}^D q_i \right) - \frac{\lambda}{\sqrt{D}}$$
$$S_{\max} = \frac{1}{D} \left(\sum_{i=1}^D q_i \right) + \frac{\lambda}{\sqrt{D}}$$



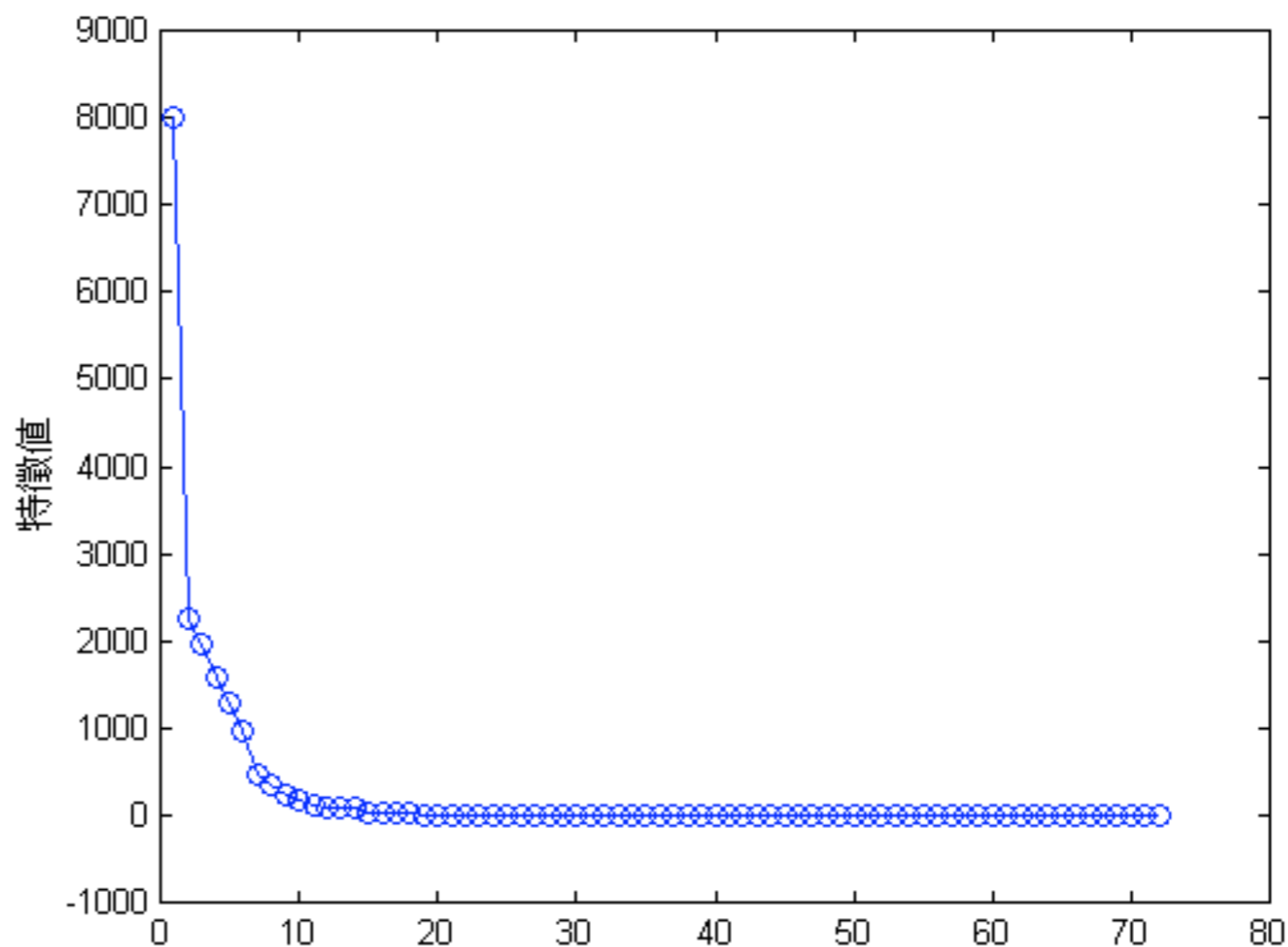


Equal-average Hyperplane Partitioning Method

- 因為我們將每個資料點平移到平均值為0，所以必須另外找一條直線 L 代替原本的中心線
 - 用 **principal component analysis** 找出 L
 - 令 M 矩陣的每一列為資料點 R_{jk} 的座標
 - 求出 $M^T M$ 的eigenvalue和eigenvector
 - 對應到最大eigenvalue的eigenvector即為所求
- 用 $\frac{R_{jk} \times L}{\sqrt{D}}$ 代替座標平均值

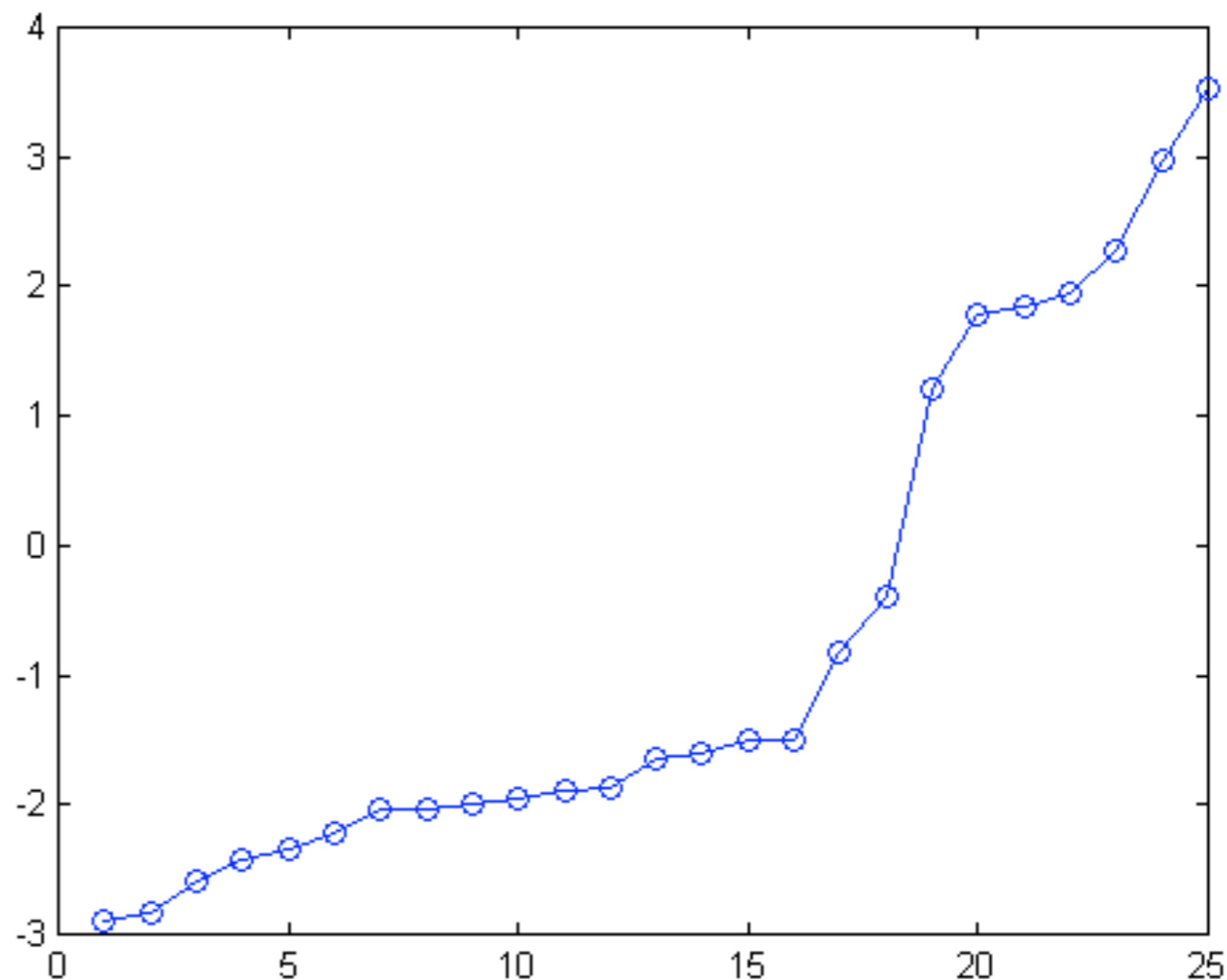
Equal-average Hyperplane Partitioning Method

- 資料庫中某一首歌曲有**25**個片段，每個片段長度為**72**
- 用**PCA**找出來的**72**個特徵值從大到小排序



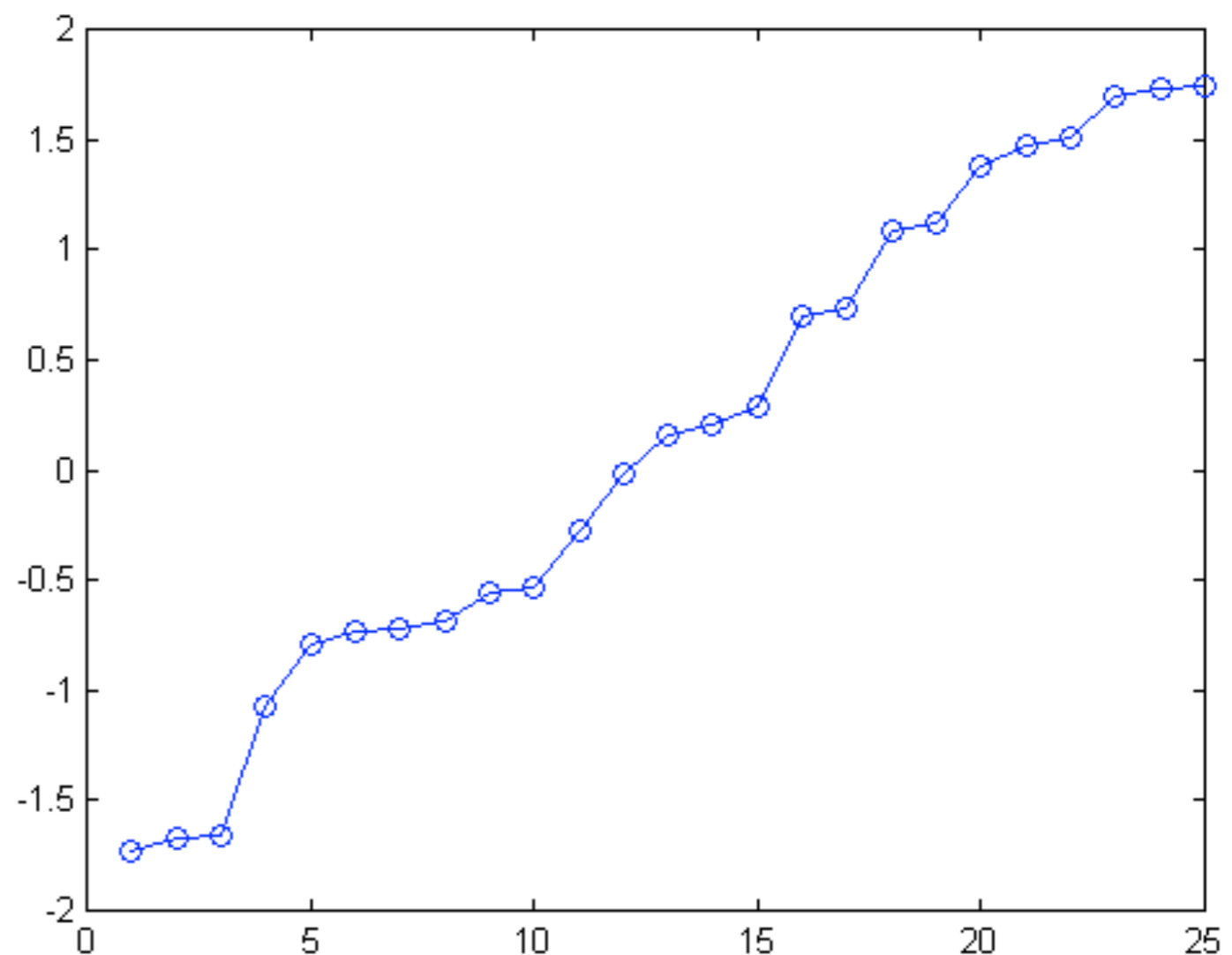
Equal-average Hyperplane Partitioning Method

- 25個片段投影在對應到最大特徵值的特徵向量後的座標平均值從小到大排序



Equal-average Hyperplane Partitioning Method

- 25個片段投影在對應到第二大特徵值的特徵向量後的座標平均值從小到大排序





實驗結果

- 實驗環境

- 電腦 P III 800 , 256MB RAM , Windows 2000

- 資料庫

- 8552首 , 包括中文、台語、英文和日文歌曲

- 測試歌聲wav檔

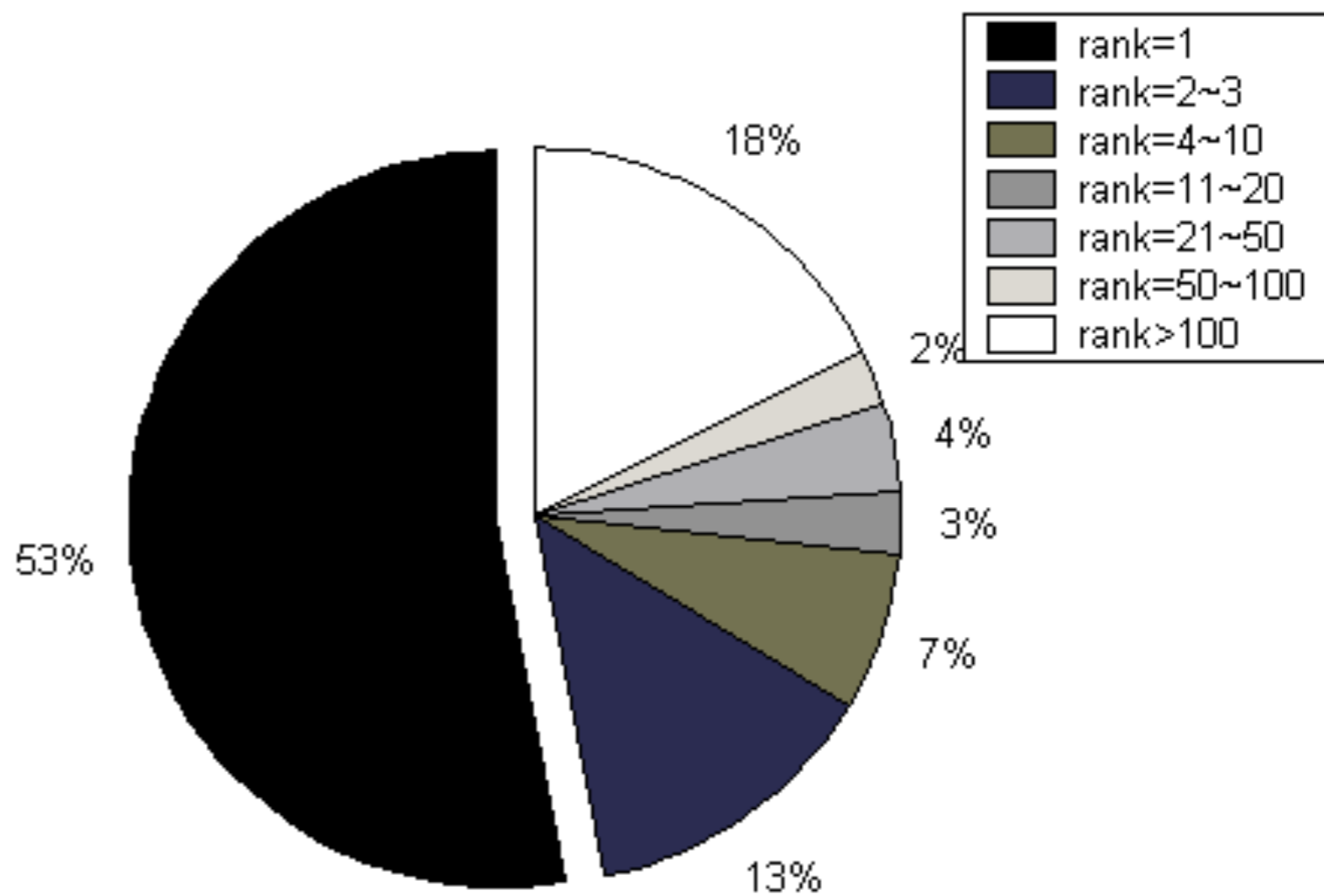
- 從頭唱 1054 首

- 從任意處唱 1650 首

- 每一首長 8 秒鐘

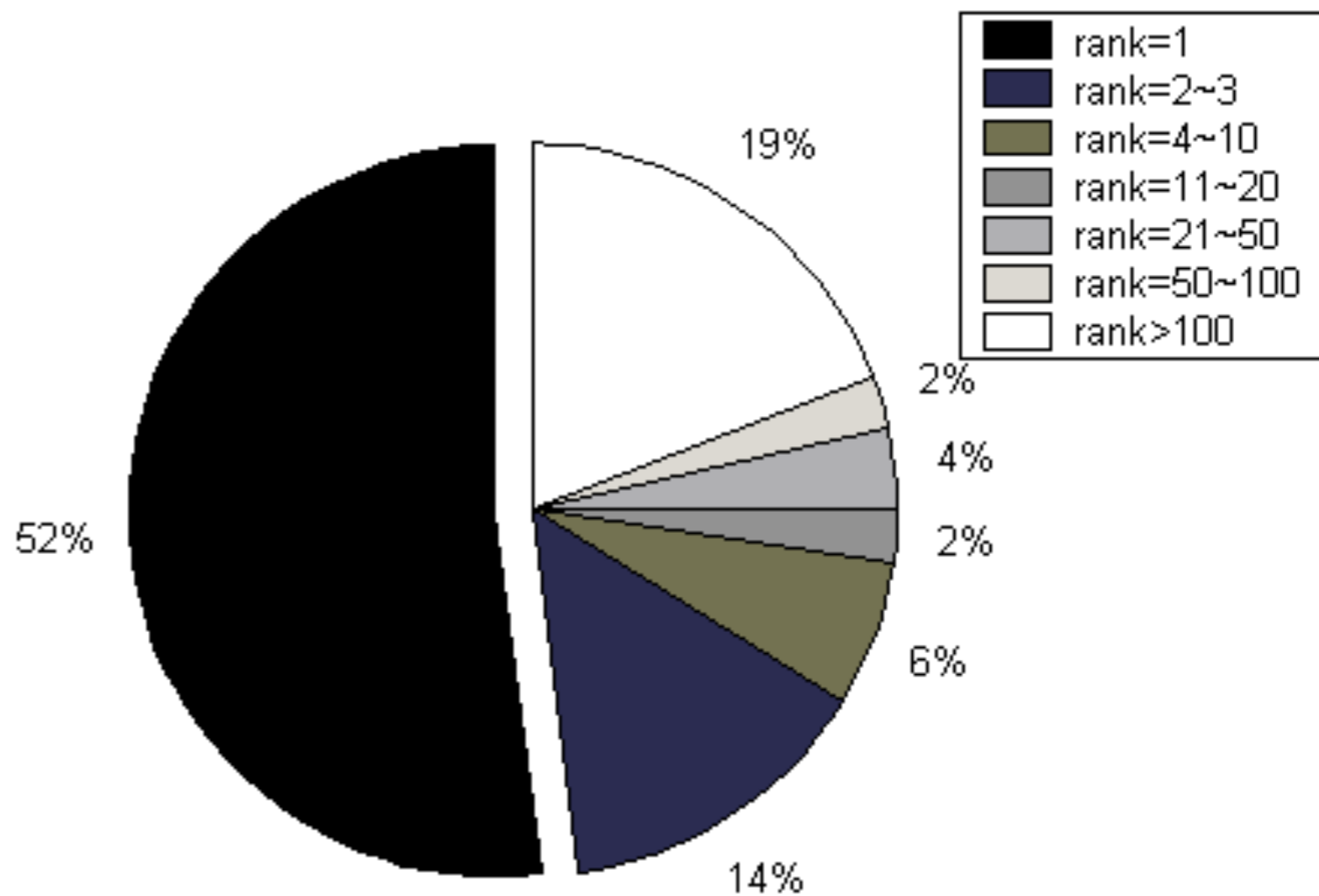
實驗結果—從頭比對

- 只用dtw
- 前三名**66.41%**
- 平均搜尋時間
16.71秒



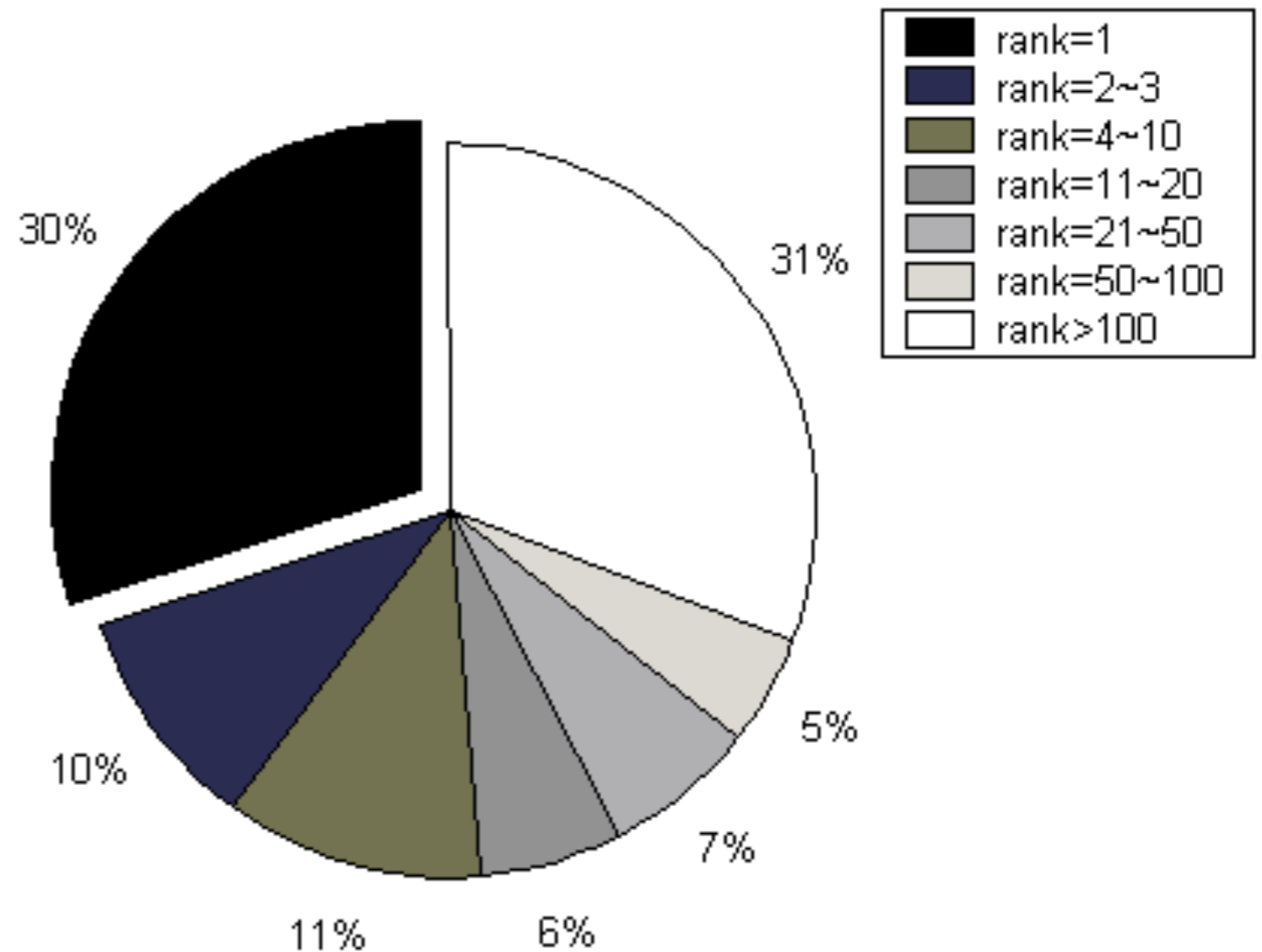
實驗結果—從頭比對

- 只用改良式dtw，保留100名距離
- 前三名66.13%
- 平均搜尋時間6.96秒



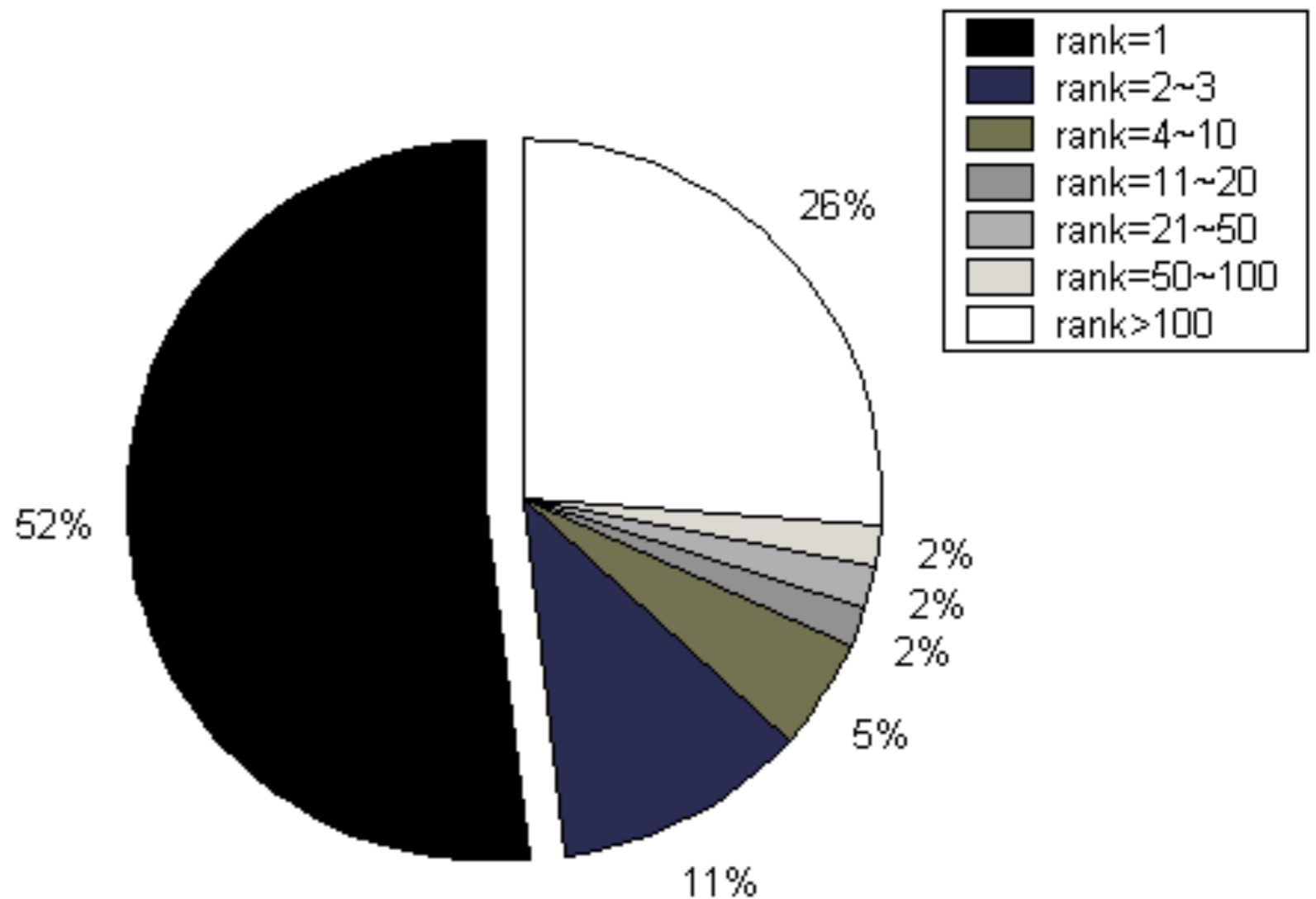
實驗結果—從頭比對

- 只用linear scaling
- 前三名39.94%
- 平均搜尋時間0.1秒

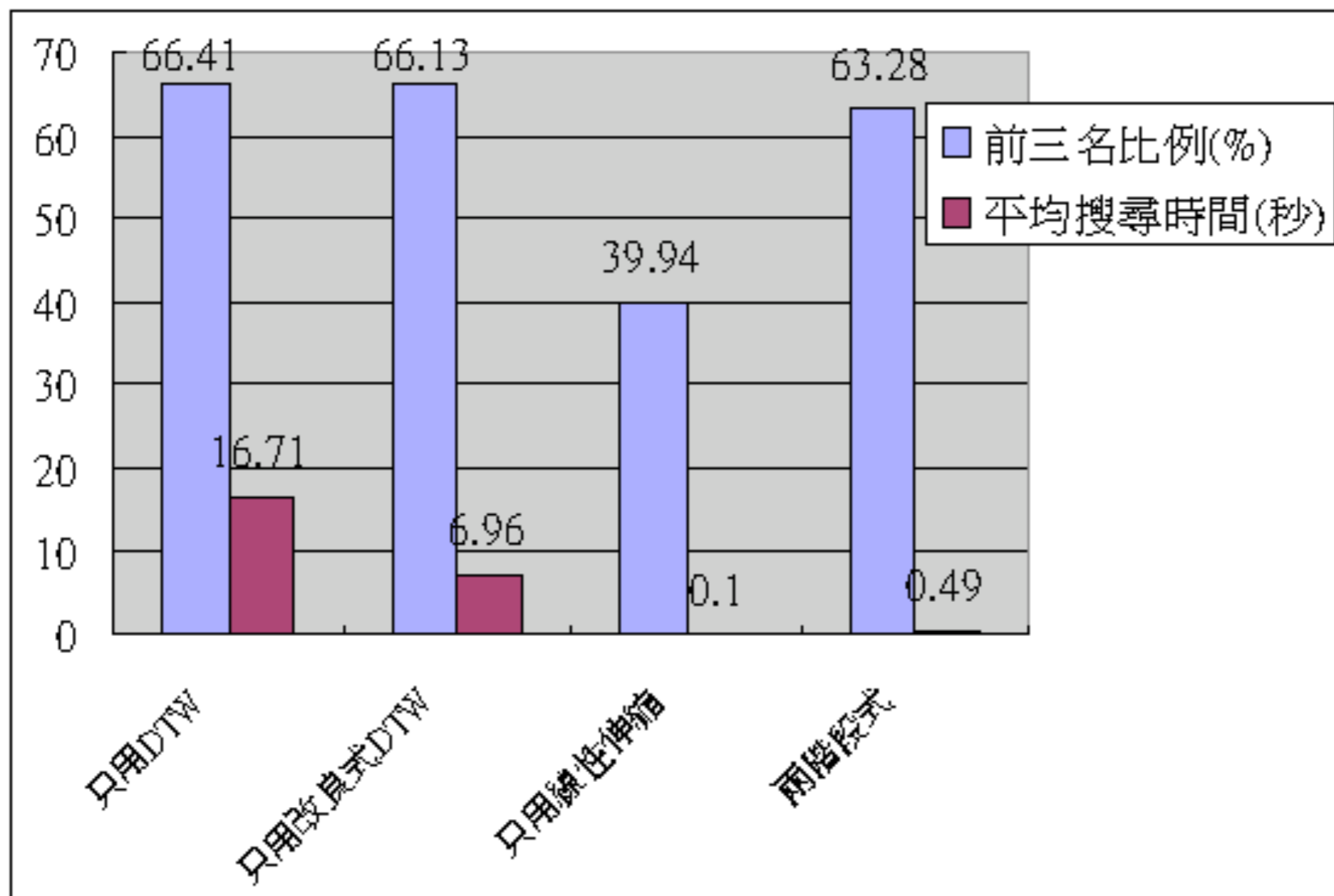


實驗結果—從頭比對

- 用兩階段式比對
- 第一階段保留**200**首歌進入第二階段
- 前三名**63.28%**
- 平均搜尋時間**0.49**秒

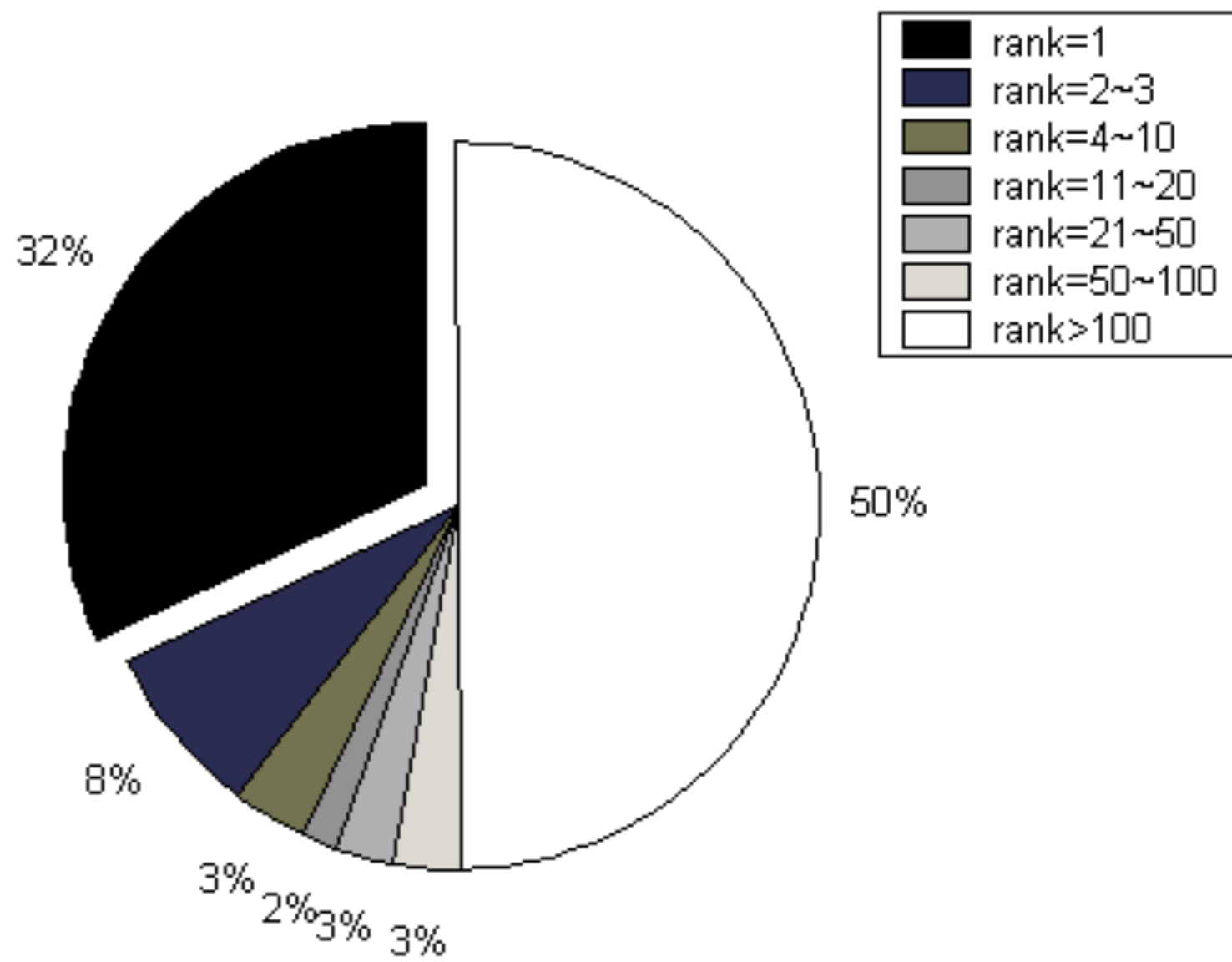


實驗結果—從頭比對比較圖

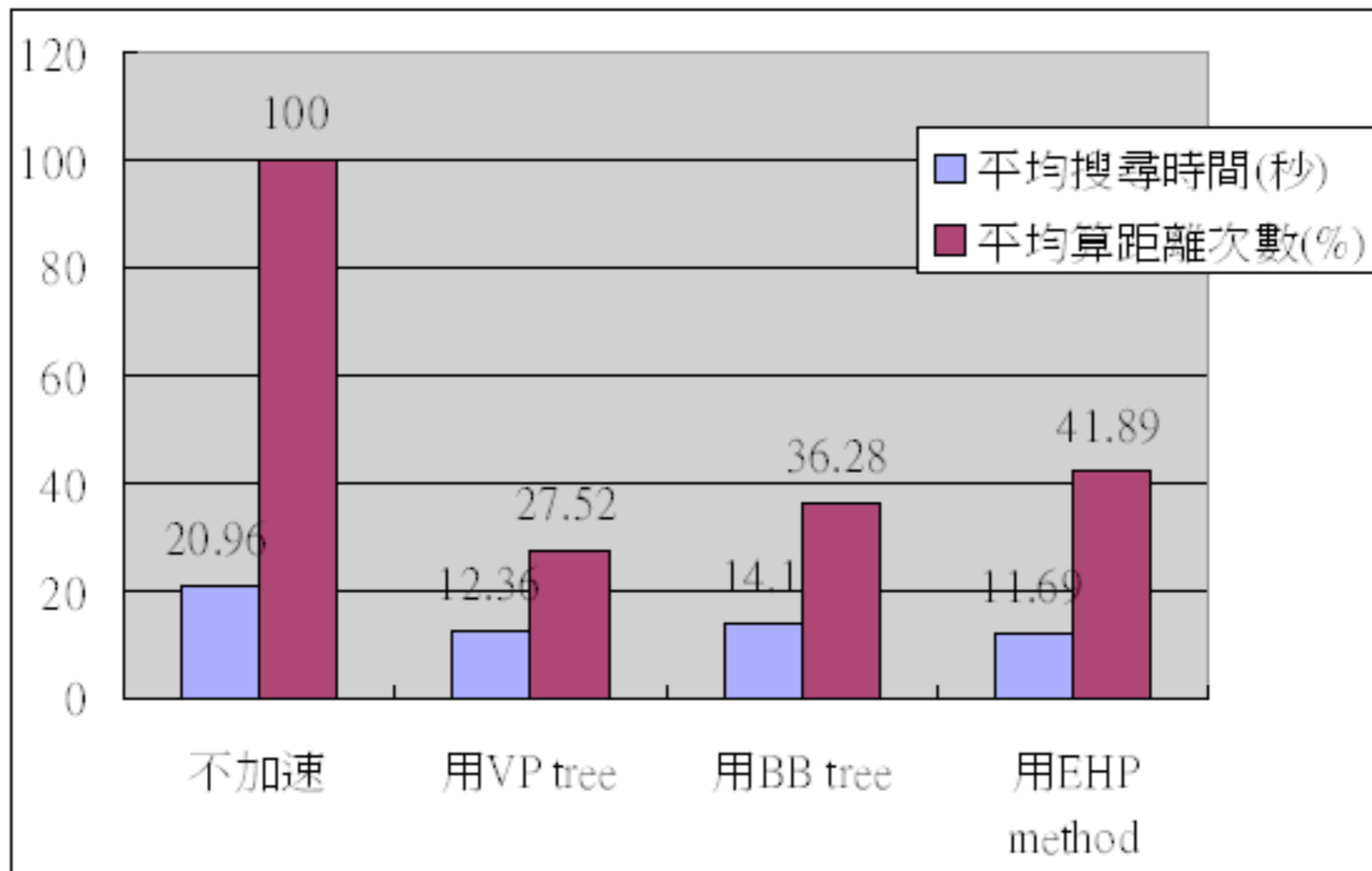


實驗結果—全曲比對

- 用兩階段式比對
- 前三名**39.63%**



實驗結果—全曲比對比較圖





錯誤分析

- 資料庫錯誤
 - 有前奏
 - 同一時間不只一個音符
 - 主旋律錯誤
 - 相同歌曲但不同歌名
- 測試歌聲不良
 - 有雜音
 - 氣不足、音不準
 - 唱錯
 - 拍子不準



結論

- 從頭比對

- 只用dtw → linear scaling + 改良式dtw

- 大約快 34 倍
- 前三名比例下降不到 4%

- 全曲比對

- 只用改良式dtw → linear scaling + 改良式dtw

- 大約快 23 倍
- 第一階段不用加速方法 → 第一階段用equal-average hyperplane partitioning method
 - 大約快 1.8 倍
 - 平均一首歌約多佔 2.7 K byte 的記憶體



3. Video retrieval techniques



Differences and relations between image and video

- Images are static, but video are dynamic.
- Video stream can be viewed as sequence of image frames.



Introduction

➤ Traditional Library search method

PONI Advanced Search - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://poni.smu.edu/cgi-bin/Pwebrecon.cgi?PAGE=bbSearch&SE

PONI Public ONLINE Information
The Southern Methodist University Libraries Catalog

Start Over
Contact the Libraries

New Search Headings Titles Search History PONI Databases Online Resources Local Libraries Interlibrary Loan My Library Account Request Help

Database Name: Southern Methodist University Libraries

Basic **Advanced** Course Reserves

Search for: video all of these Search in: Keyword Anywhere
 AND OR NOT

Search for: sound all of these Search in: Keyword Anywhere
 AND OR NOT

Search for: clip all of these Search in: Keyword Anywhere
 AND OR NOT

50 records per page Search Reset Select More Search Limits

New Search Headings Titles Search History PONI Databases Online Resources Local Libraries Interlibrary Loan My Library Account Request Help

Contact the Southern Methodist University Libraries

top

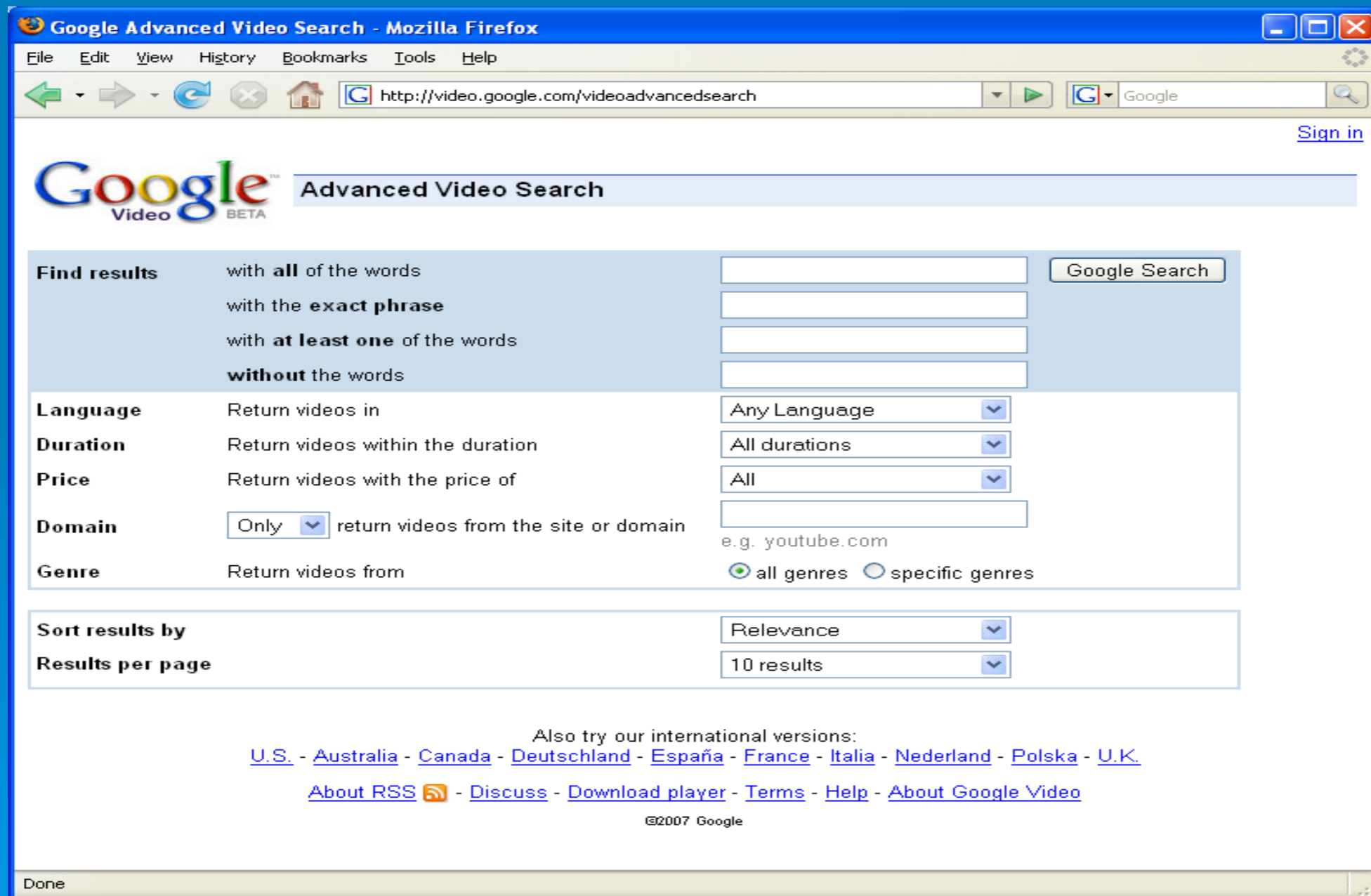
Done poni.smu.edu

Introduction (cont.)

- Other search engines still using description search method.
- Current image search method: by description.

Introduction (cont.)

➤ Sample of Google Video Search:



The screenshot shows the Google Advanced Video Search interface in Mozilla Firefox. The browser window title is "Google Advanced Video Search - Mozilla Firefox". The address bar shows the URL "http://video.google.com/videoadvancedsearch". The page features the Google logo and "Advanced Video Search" text. Below the logo, there are several search filters and options:

- Find results:** with **all** of the words, with the **exact phrase**, with **at least one** of the words, **without** the words. Each option has a corresponding text input field.
- Language:** Return videos in **Any Language** (dropdown menu).
- Duration:** Return videos within the duration **All durations** (dropdown menu).
- Price:** Return videos with the price of **All** (dropdown menu).
- Domain:** **Only** (dropdown menu) return videos from the site or domain. Input field: e.g. youtube.com.
- Genre:** Return videos from **all genres** **specific genres**.
- Sort results by:** **Relevance** (dropdown menu).
- Results per page:** **10 results** (dropdown menu).

At the bottom, there are links for international versions: [U.S.](#) - [Australia](#) - [Canada](#) - [Deutschland](#) - [España](#) - [France](#) - [Italia](#) - [Nederland](#) - [Polska](#) - [U.K.](#). Below these are links for [About RSS](#), [Discuss](#), [Download player](#), [Terms](#), [Help](#), and [About Google Video](#). The footer shows "©2007 Google".

Introduction (cont.)

➤ Google Video Archive selections:

The screenshot shows the Google Video website in a Mozilla Firefox browser window. The browser's address bar displays the URL: <http://video.google.com/?ie=UTF-8&oe=UTF-8&hl=en&q=&tab=iv>. The page features the Google Video logo with a 'BETA' tag and a search bar. Navigation tabs include 'Web', 'Images', 'Video', 'News', 'Maps', and 'more'. Below the search bar are category buttons: 'Top 100', 'Comedy', 'Music videos', 'Sports', 'Animation', 'TV shows', and 'Google Picks'. The main content area is divided into several sections:

- Popular**: A carousel of video thumbnails with titles, star ratings, and durations. Visible titles include 'white girls walk it out' (3 min), 'really expensive cat toy' (2 min), 'Bhaktha_Kannappa01.wmv' (1 hr 15 min), 'Drowning Pool - "Bodies"' (4 min), 'Baby Panda Sneeze.wmv' (13 sec), and 'Melissa Theuriau' (8 min).
- New Features**: A section with a YouTube logo and the text: 'New! Now you can also search for videos from YouTube.'
- Blog Buzz**: A list of 10 blog entries with titles such as 'Hearing on Tillman, Lynch Incidents: Jessica Lynch's Opening', 'The Simpsons Marge On Google', and 'Portugal, Lisboa. Revolução de 25 de Abril de 1974'.
- Featured**: A carousel of featured video thumbnails. Visible titles include '£10 000 coin domino effect' (2 min), 'LEGO Pirates: Treasure' (3 min), and 'Penguins go for a stroll' (1 min).
- Movers & Shakers**: A section at the bottom right of the main content area.

The browser's status bar at the bottom left shows 'Done'.

Introduction (cont.)

- Picture is worth a thousand words.
- More than words can express.
- Growing number video clips on MySpace and YouTube, there is a need for a video search engine.

Introduction (cont.)

➤ Sample YouTube Video page:

The screenshot shows the YouTube homepage in a Mozilla Firefox browser window. The address bar displays the URL <http://www.youtube.com/browse?s=mp>. The page features the YouTube logo and navigation tabs for Videos, Categories, Channels, and Community. A search bar is located in the top right corner. The main content area is titled "Most Viewed (Today)" and displays a grid of video thumbnails with titles, durations, and ratings. On the left side, there are filters for Browse, Time, Category, and Language. On the right side, there are advertisements for RealAge, AskMen, and Zune.

YouTube - Broadcast Yourself. - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.youtube.com/browse?s=mp

Sign Up | My Account | History | Help | Log In

Search

Videos Categories Channels Community Upload Videos

Videos Most Viewed (Today) My Videos - Favorites - Playlists - Inbox - Subscriptions

Browse

- Most Recent
- Most Viewed
- Top Rated
- Most Discussed
- Top Favorites
- Most Linked
- Recently Featured
- Most Responed
- Watch on Mobile

Time

- Today
- This Week
- This Month
- All Time

Category

- All
- Autos & Vehicles
- Comedy
- Entertainment
- Film & Animation
- Gadgets & Games
- Howto & DIY
- Music
- News & Politics
- People & Blogs
- Pets & Animals
- Sports
- Travel & Places

Language

- All
- English
- Spanish
- Japanese
- German
- Chinese

Pages: 1 2 3 4 5 Next

Ads by Goooooogle

The Most Amazing Basketball Shot EVER
00:17
Added: 1 day ago
From: [kasnawalo](#)
Views: 259,750
★★★★★
763 ratings

15 year old tries to out drive the police at 150MPH!
04:45
Added: 1 day ago
From: [kasnawalo](#)
Views: 189,132
★★★★★
765 ratings

Don't Play Entropia Universe
00:17
Added: 12 hours ago
From: [EUMWarning](#)
Views: 153,645
★★★★★
100 ratings

LINK LIVEFOOTSTREAM S CHARGES £5 FOR FREEHERE NO CHARGE!
03:12
Added: 1 day ago
From: [AndAgain432](#)
Views: 117,313
★★★★★
116 ratings

Prepare to be Shocked
You may be younger than you think. Take the RealAge test and find out.
[www.RealAge.com](#)

Weight Loss For Men
The top 10 things guys need to know about fitness. Free tips at AskMen.
[www.AskMen.com](#)

Exhausted All The Time ?
It's Not Your Fault. You Just Need To Boost Your HGH Levels. It's Easy
[www.HGH-Facts.com](#)

Copy your iPod music
Transfer songs from your iPod or iPod Mini to your computer.
[www.findleydesigns.com](#)

ZUNE Unlimited Downloads
Music, Movies, MP4 Videos, Sports Instant Access for Only \$38.96
[www.ZuneReactor.com](#)

4 year old child injured at Colorado state football practice
00:34
Added: 1 day ago
From: [kimmelscorner](#)
Views: 108,910
★★★★★
161 ratings

Order of the Phoenix Trailer No. 2
02:12
Added: 1 day ago
From: [leakynewsdotcom](#)
Views: 102,481
★★★★★
395 ratings

Top 10 Tips to Get Bathing Suit Ready: sparkpeople.com top10
02:11
Added: 1 day ago
From: [elfersp](#)
Views: 100,358
★★★★★
88 ratings

Harry Potter and the Order of the Phoenix Domestic Trailer
02:13
Added: 1 day ago
From: [mnnvideos](#)
Views: 99,230
★★★★★
179 ratings

LIVEFOOTSTREAM

Scarlett Wipes Out

Will Ferrell's landlord

Diet.com Weight

Done

Introduction (cont.)

- Therefore, we need a better search technique – Content-Based Video Retrieval System (CBVR).

Image Retrieval Phase

- Query by example (QBE)
 - Allow to select sample image to search.

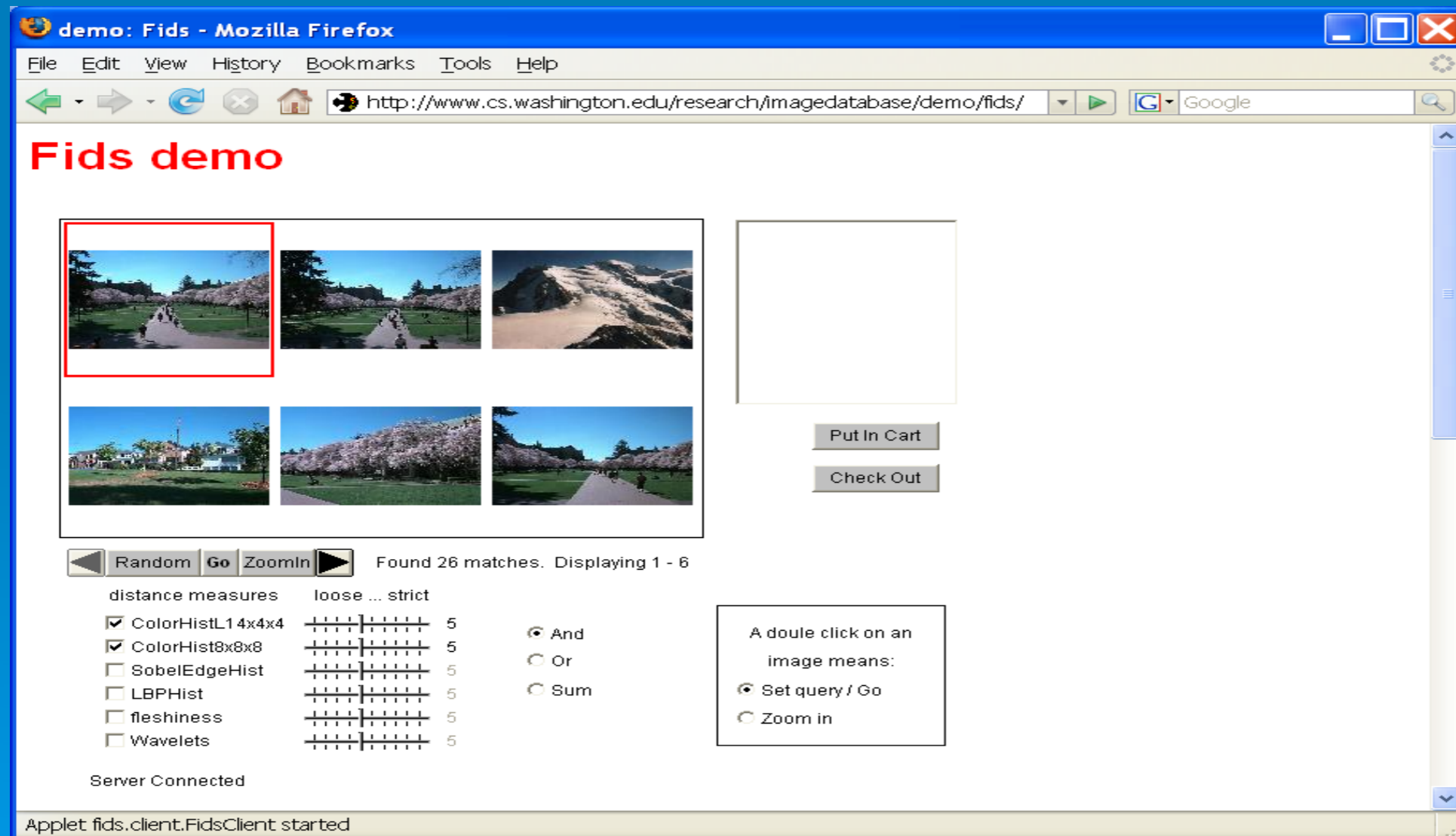


Image Retrieval Phase (cont.)

Yet Another CBVR Application Interface



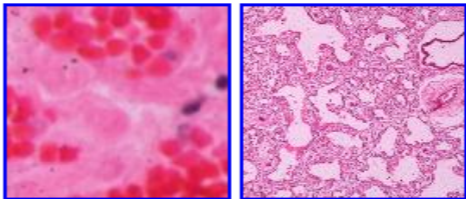
Content-Based Image Retrieval - Mozilla Firefox


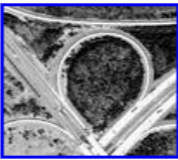
File Edit View History Bookmarks Tools Help

http://wang14.ist.psu.edu/cgi-bin/zwang/regionsearch_show.cgi

Google

D e m o

Main Image Class ->	Photographs	Graphs	Biomedicine
Option 1	Click <i>Random</i>	Click <i>Random</i>	Click <i>Random</i>
Option 2	Query image URL or ID <input type="text"/>	Query image URL or ID <input type="text"/>	
Option 3	java drawing interface (not linked to db)	java drawing interface (not linked to db)	
Option 4	Start with 	Start with 	Start with 

Main Image Class ->	Art - Cultural	Satellite	Paintings
Option 1	Click <i>Random</i>	Click <i>Random</i>	Click <i>Random</i>
Option 2	Start with 	Start with 	Start with 

Done

Introduction (cont.)

- What good is video retrieval?
 - Historical Achieve
 - Forensic documents
 - Fingerprint & DNA matching
 - Security usage

Overview (cont.)

➤ CBVR has two Approaches:

- Attribute based
- Object based

➤ CBVR can be done by:

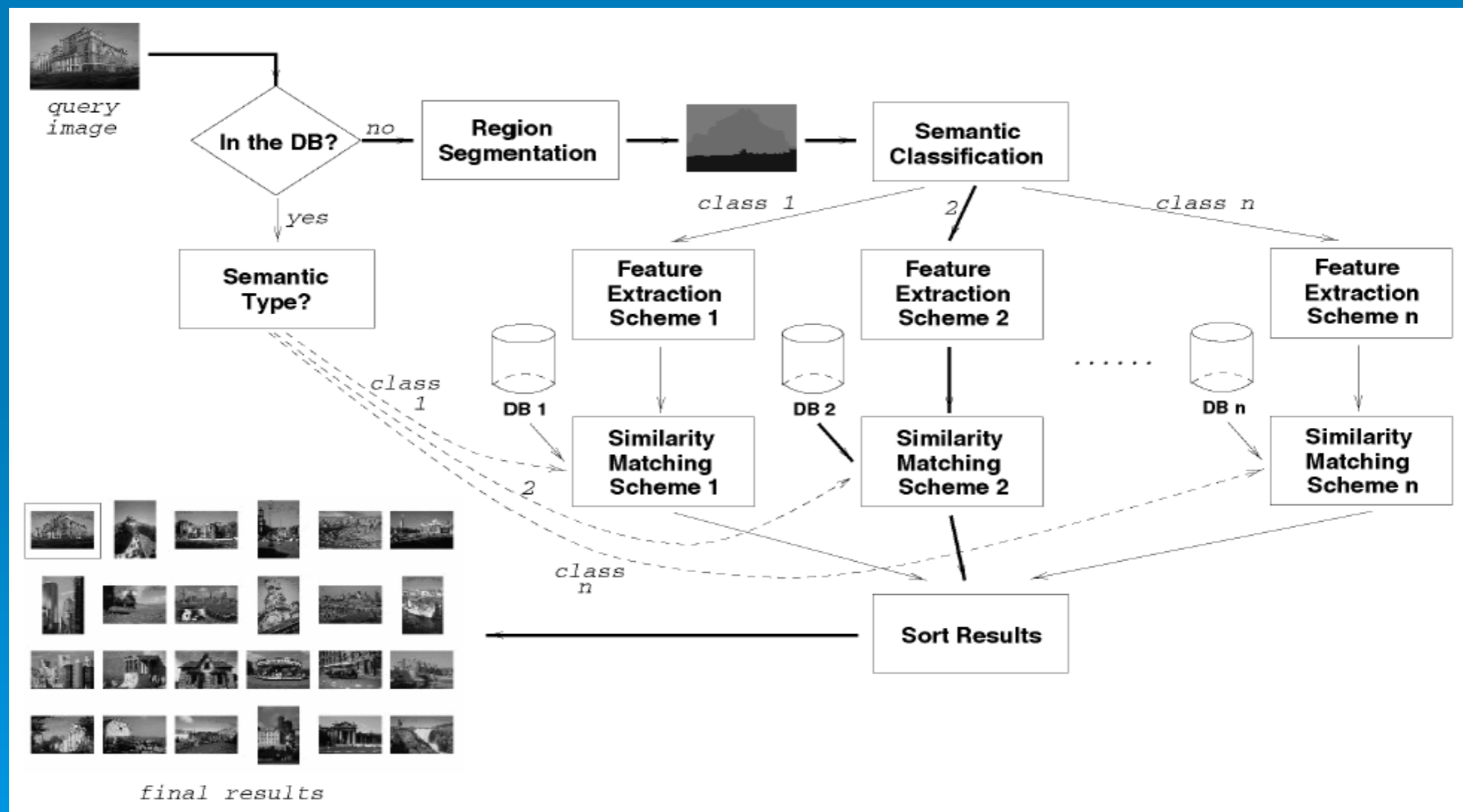
- Color
- Texture
- Shape
- Spatial relationship
- Semantic primitives
- Browsing
- Objective Attribute
- Subjective Attribute
- Motion
- Text & domain concepts

Overview (cont.)

- CBVR has two phases:
 - Database Population phase
 - Video shot boundary detection
 - Key Frames selection
 - Feature extraction
 - Video Retrieval phase
 - Similarity measure

Overview (cont.)

➤ How CBVR works:



Database Population Phase

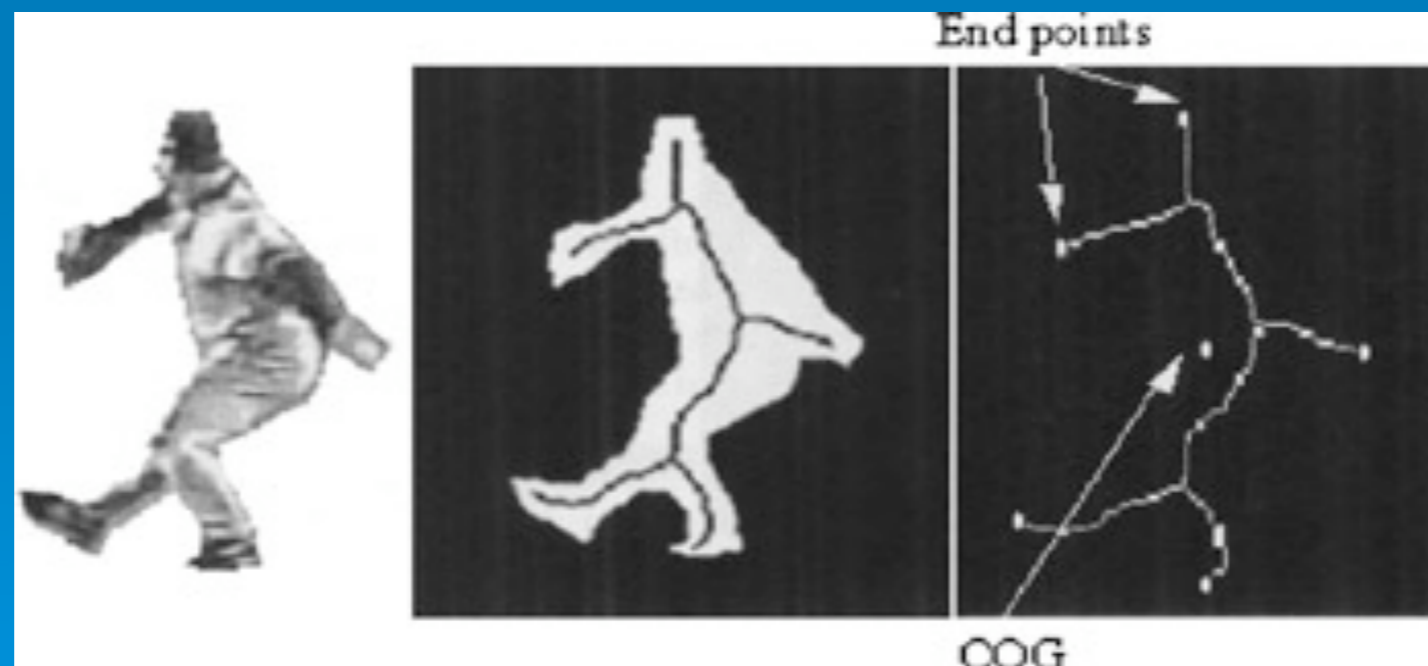
- Here are the three major procedures:
 - Shot boundary detection – partition, segments



[Luo, Hwang, Wu, 2004]

Database Population Phase (cont)

- Key frames selection – select characteristics
- Extracting low-level spatial features like color, texture, shape, etc.



[Luo, Hwang, Wu, 2004]

Database Population Phase (cont.)

- Video is complex data type – audio & video
- Audio can be handled by query by humming.
- Voice recognition system using Patricia-like tree to construct all possible substrings of a sentence.
- Audio is categorized by: speech, music, and sound.
- Audio retrieval methods: Hidden Markov Model, Boolean Search with multi-query using Fuzzy Logic.

Database Population Phase (cont)

- Most simple database storage: description of video as index along with the video.
- Human effort is involved in this case.
- We are searching for automatic video indexing and digital image storage method – Latent Semantic Indexing (LSI)

Database Population Phase (cont.)

- LSI is using vector space model – low rank approximation of vector space represent image document collection.
- Original matrix is replaced by an as close as possible matrix, where its column space is only the subspace of the original matrix column space.
- By reducing the rank of the matrix, noises (duplicate frames) are reduce to improve storage and retrieval performance.
- Term indexing is referred to the process of assigning terms to the content of the video.

Database Population Phase (cont.)

- Closest terms in the database is returned based on the similarity measure between the query images and the resulting ones.
- Cosine similarity measure is used in the vector space model.
- Cosine similarity measure on Term-by-video matrix:

$$\cos(t_1, v) = \frac{\sum_{h=1}^k t_1 v}{\sqrt{\sum_{h=1}^k t_1^2 \sum_{h=1}^k v^2}}$$

Database Population Phase (cont.)

- Enterprise database like Oracle introduces new object type: `ORDImage`, which contains four different visual attributes: global color, local color, texture and shape.
- `ORDImageIndex` provides multidimensional index structure to speed up stored feature vectors.

Database Population Phase (cont.)

- Oracle example of joining two images of Picture1 and Picture2:

```
CREATE TABLE Picture1(  
    author VARCHAR2(30),  
    description VARCHAR2(200),  
    photo1 ORDSYS.ORDImage,  
    photo1_sig ORDSYS.ORDImageSignature  
);  
  
CREATE TABLE Picture2(  
    mydescription VARCHAR2(200),  
    photo2 ORDSYS.ORDImage,  
    photo2_sig ORDSYS.ORDImageSignature  
);  
  
SELECT p1.description, p2.mydescription  
FROM Picture p1, Picture p2,  
WHERE  
  
        ORDSYS.IMGSimilar(p1.photo1_sig, p2.photo2_sig,  
        'color="0,6" texture="0,2" shape="0,1"  
        location="0,1"', 20)=1;
```

Note: Weighted sum of the distance of the visual attributes is less than or equal to the threshold, the image is matched.

Homework

- Implement a naïve image retrieval system
 - Using color moments

