

**MPEG-7 Audio:  
The MelodyContour Description Scheme**

**Music Retrieval by Melodic Query**

Youngmoo Kim  
Machine Listening Group  
MIT Media Lab



# Application

---

- “Name that tune” for computers
  - Melodies are an intuitive way of relating to musical content.
  - People tend to remember melodies.
- Problem: computers are very literal (and people tend not to sing well)
  - Need a representation that works the way humans deal with melody.



# Goals of Melody Description

---

- Efficiency
  - Able to represent distinct melodies compactly
- Scalability
  - Works for both small music collections and large databases
- Robustness
  - Sung/hummed queries can be imprecise



# Elements of melody

---

- Series of related pitches
  - Invariant to transposition
  - Invariant to instrumentation
- Rhythm is also important



"Bridal Chorus" from Lohengrin  
("Here Comes the Bride")



"O Tannenbaum"



# Approach

---

- Pitch representation
  - 5-level interval contour (-2,-1,0,+1,+2)
    - Invariant to transposition
    - Robust to imprecise queries
    - Experimental and perceptual validation
    - Intervals  $>M2$  are represented by  $\pm 2$
    - Intervals of  $m2$  and  $M2$  are  $\pm 1$
- Rhythmic representation
  - Meter and nearest-beat information
    - Resolves some ambiguities resulting from using only pitch/interval information
- Automatic extraction?
  - Limited





# MPEG-7 DDL for "Moon River"

---

```
<!-- MelodyContourDS description of "Moon River" -->
<!-- (7 intervals = 8 notes total) -->
<Contour>
  <ContourData>2 -1 -1 -1 -1 -1 1</ContourData>
</Contour>
<!-- Meter of melody -->
<Meter>
  <Numerator>3</Numerator>
  <Denominator>4</Denominator>
</Meter>
<!-- Beat positions of notes -->
<!-- (8 notes = 1 more than number of intervals) -->
<Beat>
  <BeatData>1 4 5 7 8 9 9 10</BeatData>
</Beat>
```



# System block diagram

