

# Clustering Billions of Images with Large Scale Nearest Neighbor Search

Ting Liu, Charles Rosenberg, Henry A. Rowley

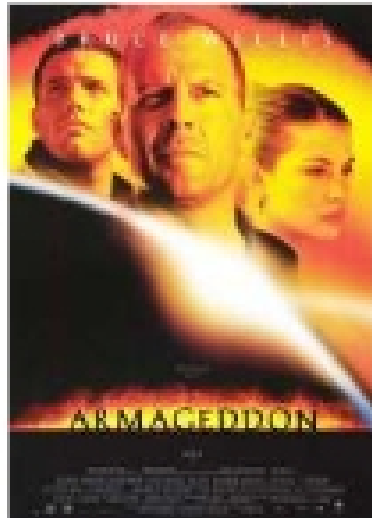
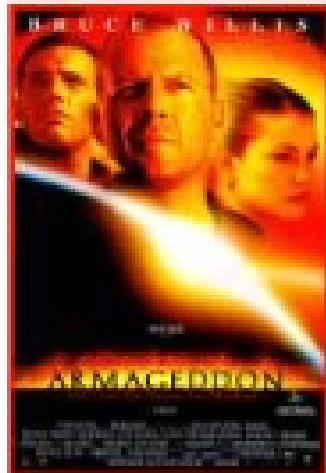
IEEE Workshop on Applications of Computer Vision  
February 2007

Presented by Dafna Bitton on May 6<sup>th</sup>, 2008  
for CSE 291

# Problem Statement

Goal 1: Find **approximate** nearest neighbors for a repository of over one billion images

Goal 2: Perform clustering based on the results



# Context of the Task

- Billions of images on the web
- Modern image search is text-based, largely due to so many images!
- Scale makes most computer vision tasks infeasible in real time