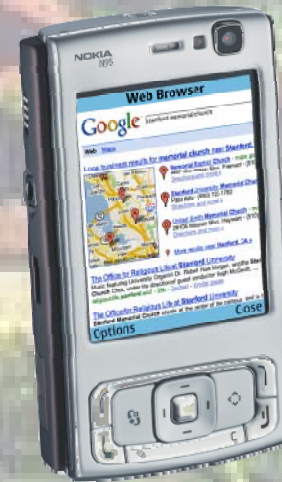


Mobile Visual Search



Bernd Girod
Stanford University
bgirod@stanford.edu



Mobile Visual Search



Mobile Visual Search

Visual Bookmarks for Movies and TV Shows

Mobile Visual Search Applications



Ads/Catalogs



Wine Labels



Museum Guide

The Amazon app includes **Amazon Remembers**

Use Amazon Remembers to create visual lists of things you want to remember while out and about. Photos you take from the app are stored on both the Amazon App and the Amazon.com site as reminders.

If the item you want to remember is a product, Amazon will try to find an item for sale like the one in the photo. If we do, we'll send you an e-mail alert and post the result along with the original photo.

- See something you want to remember**
When you notice an item to remember, tap "Remembers" in the Amazon App.
- Snap photo & send**
Your iPhone camera will open. Take a photo of the item and it will be sent to Amazon.
- See reminders**
Your photos & any similar products that Amazon finds are stored in the app and on Amazon.com.

Comparison Shopping

MOBILE IMAGE RECOGNITION TRY IT OUT NOW!!!

1. POINT YOUR MOBILE PHONE CAMERA TO THE MOVIE POSTER.

2. SNAP A PICTURE AND SEND IT!

IN SWITZERLAND: MMS TO 5555 (OR 079 394 57 00 FOR ORANGE CUSTOMERS)

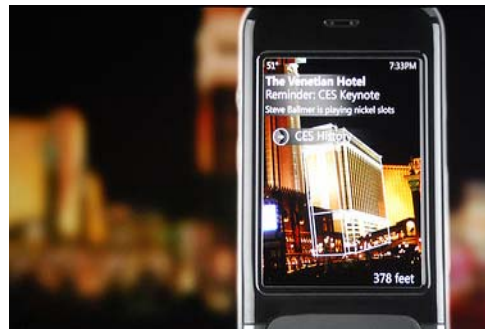
IN GERMANY: MMS TO 81000

EVERYWHERE: EMAIL TO M@KOOBARA.COM

3. FIND ALL RELEVANT INFORMATION ABOUT THE MOVIE ON YOUR MOBILE PHONE.

Show another poster Movie data provided by **metacritic**

Movie Posters



Landmarks



Real Estate

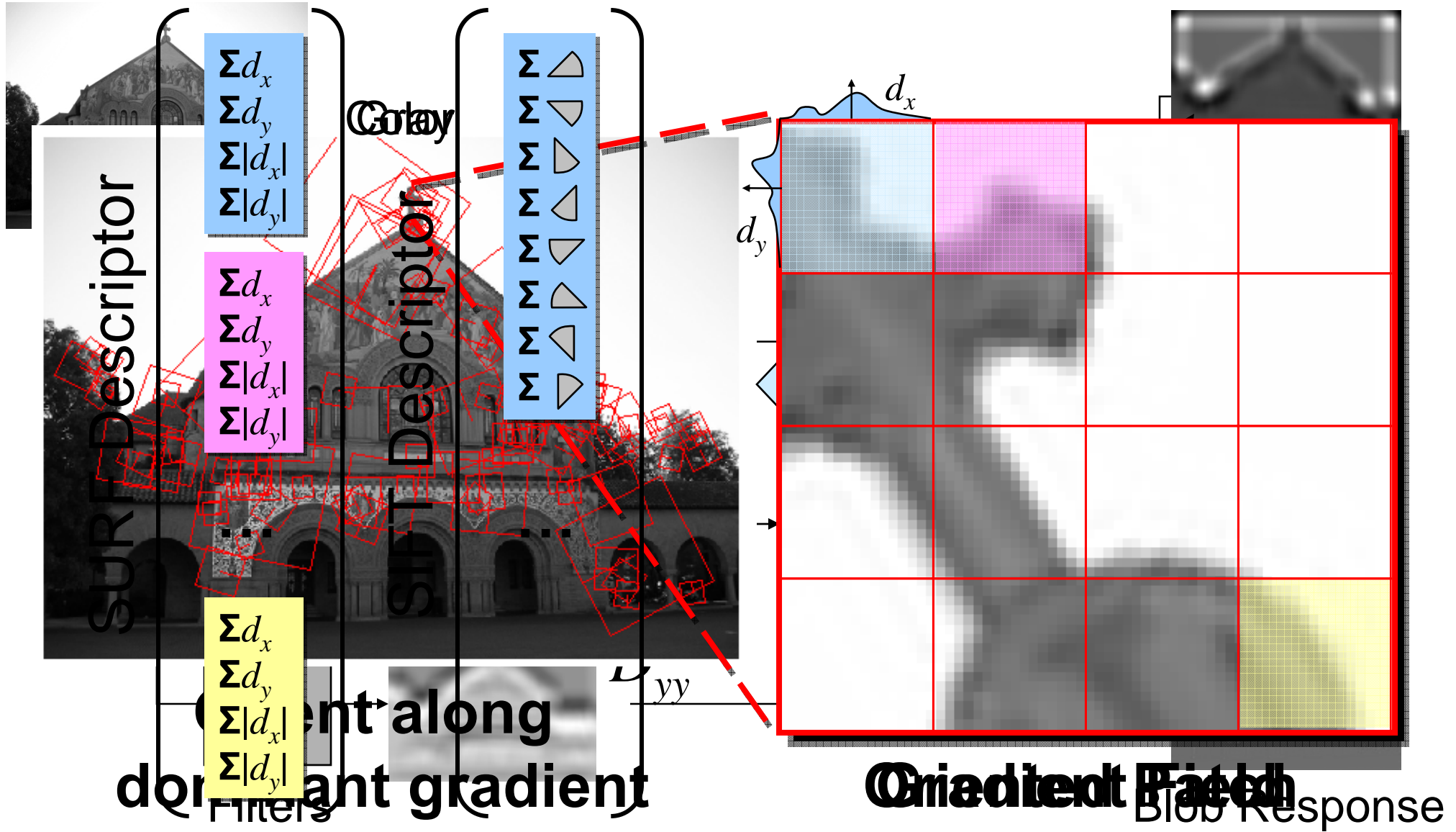


Media covers (CD, DVD, books)

Outline

- Computer vision: “Bag of Words” matching
- Feature compression
- Phone-based vs. server-based processing
- Demos (during the break)

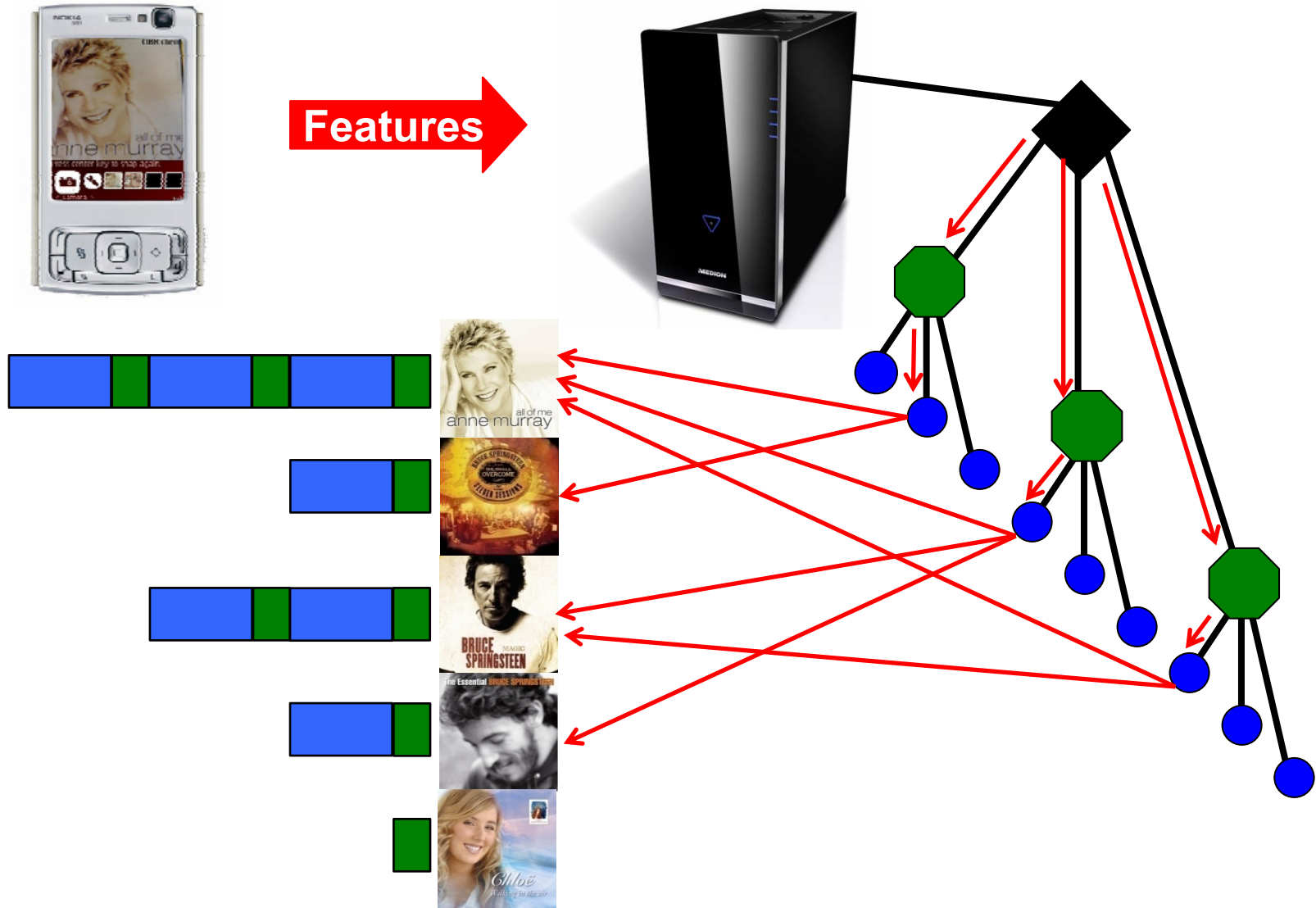
Computing Visual Words



“Bag of Words” Matching & Geometric Consistency Check



Vocabulary Tree



Architecture A: Send Image

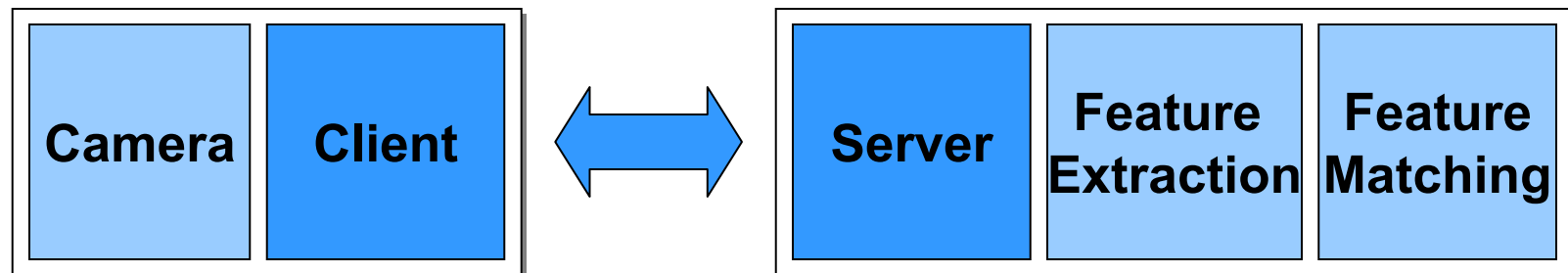


20 kbps → 20 sec

Image →



← *Information*



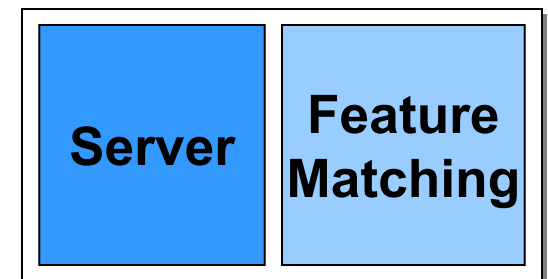
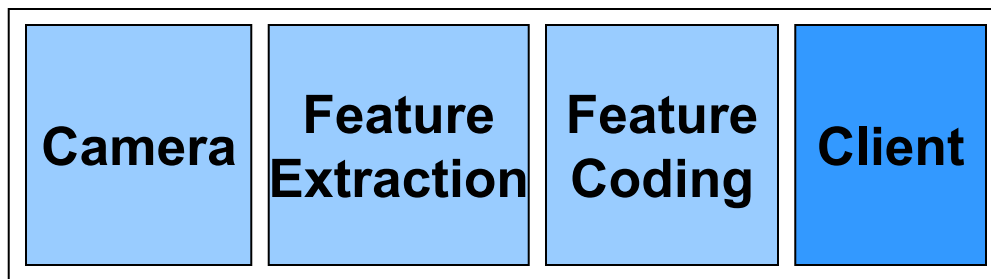
Architecture B: Send Features



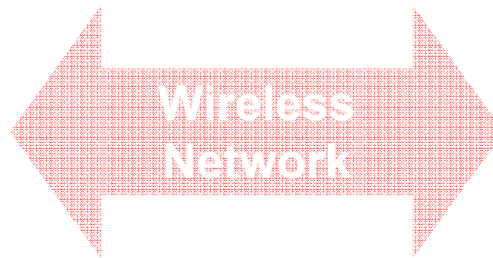
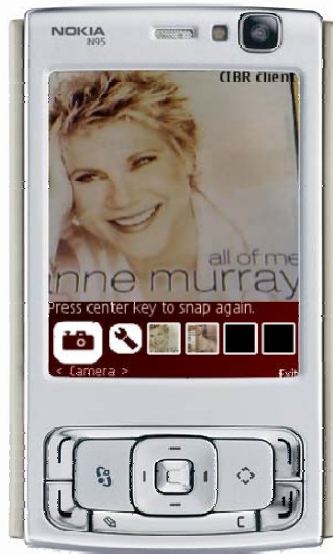
Features →



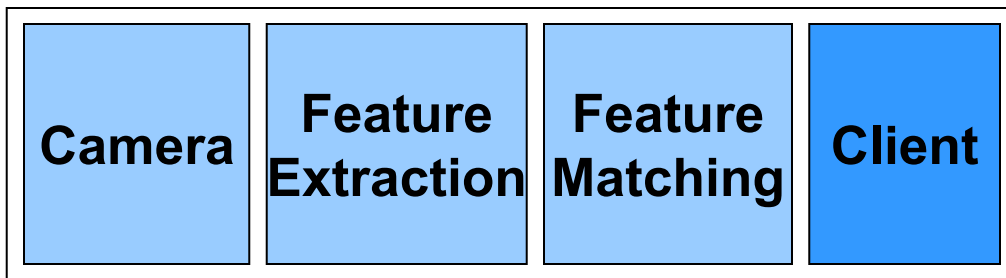
← *Information*



Architecture C: Features on Phone



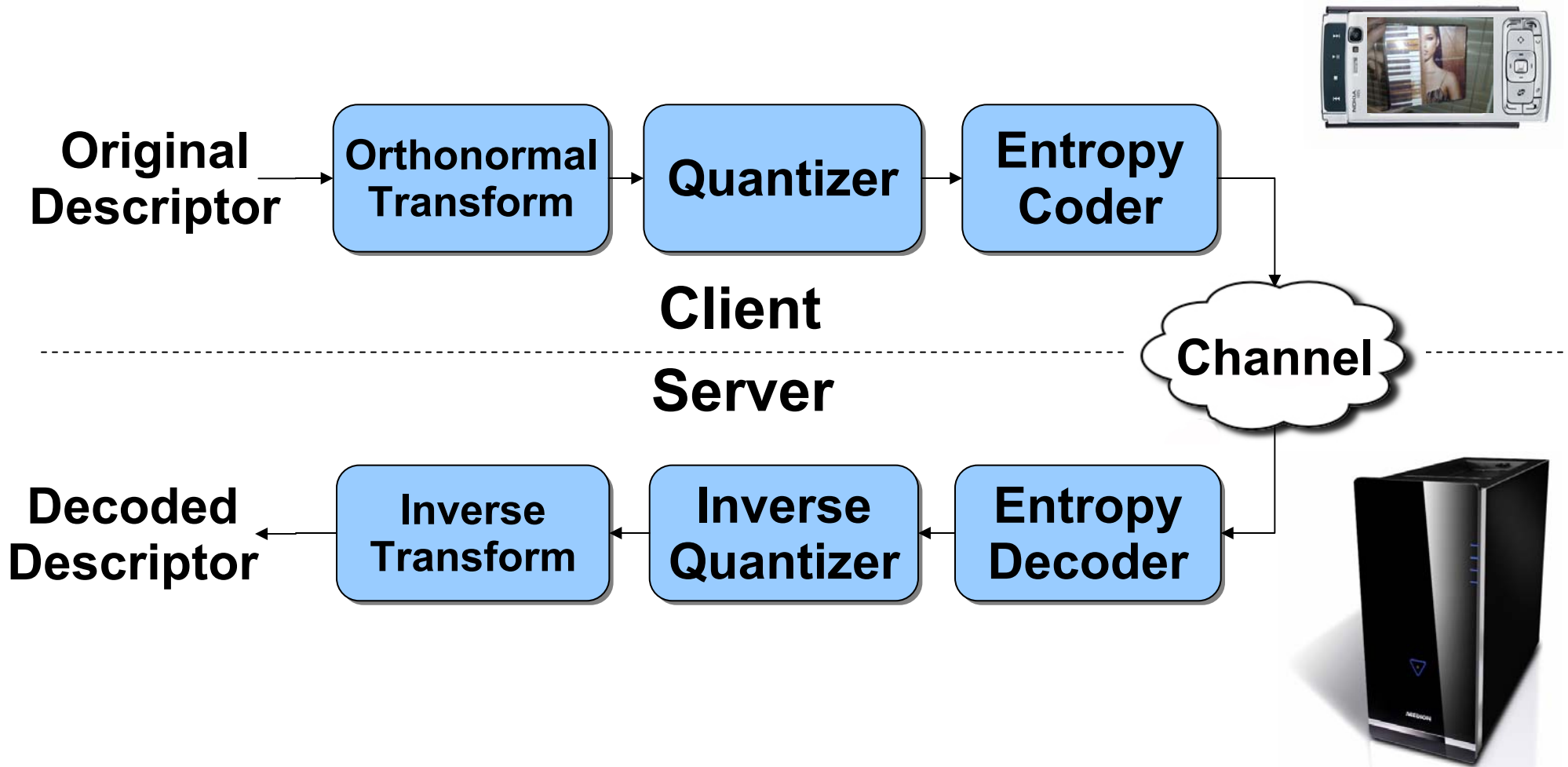
← *Features*
← *Information*



GPS-Aided Landmark Recognition

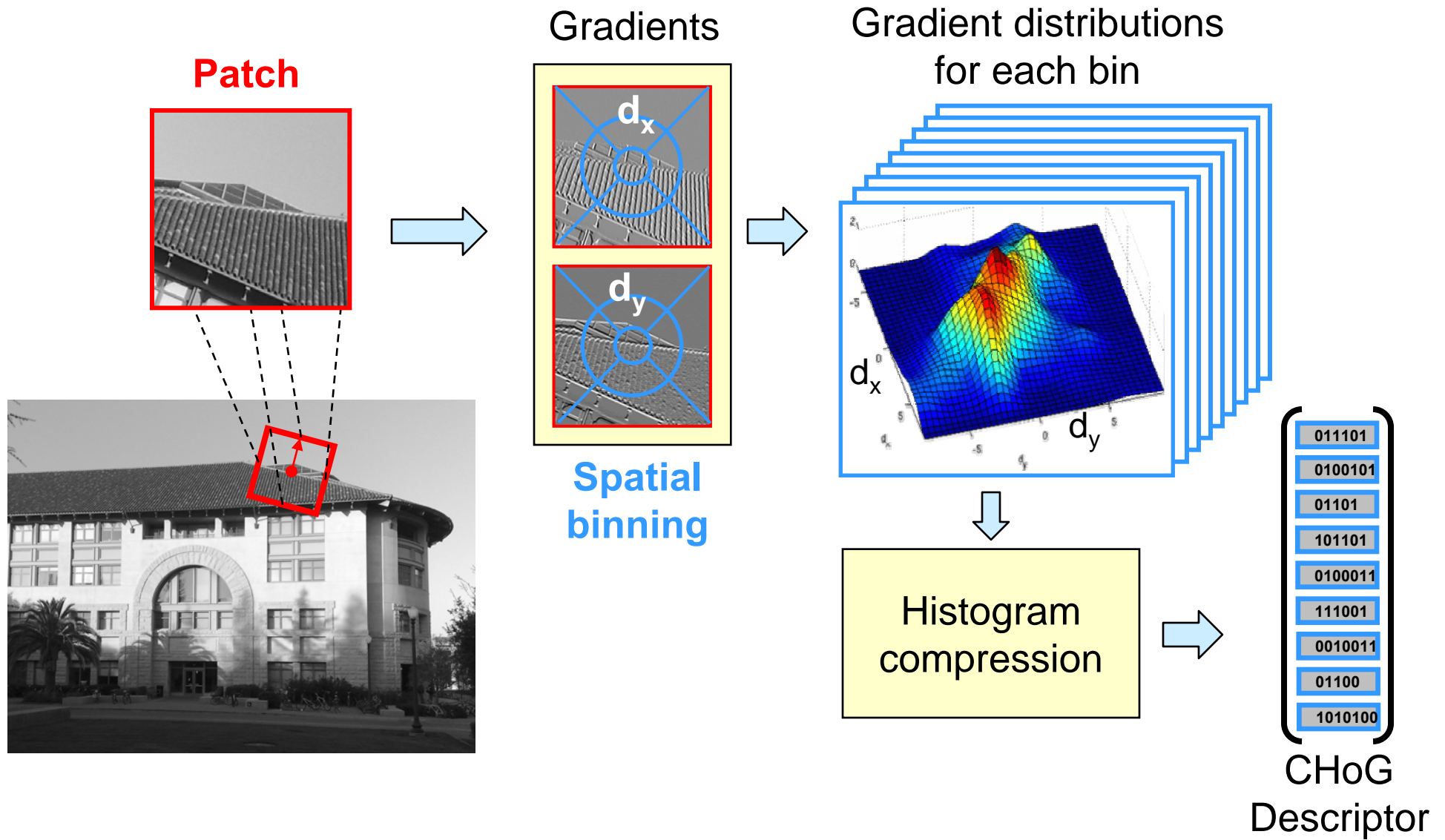


Transform Coding of SURF/SIFT Descriptors

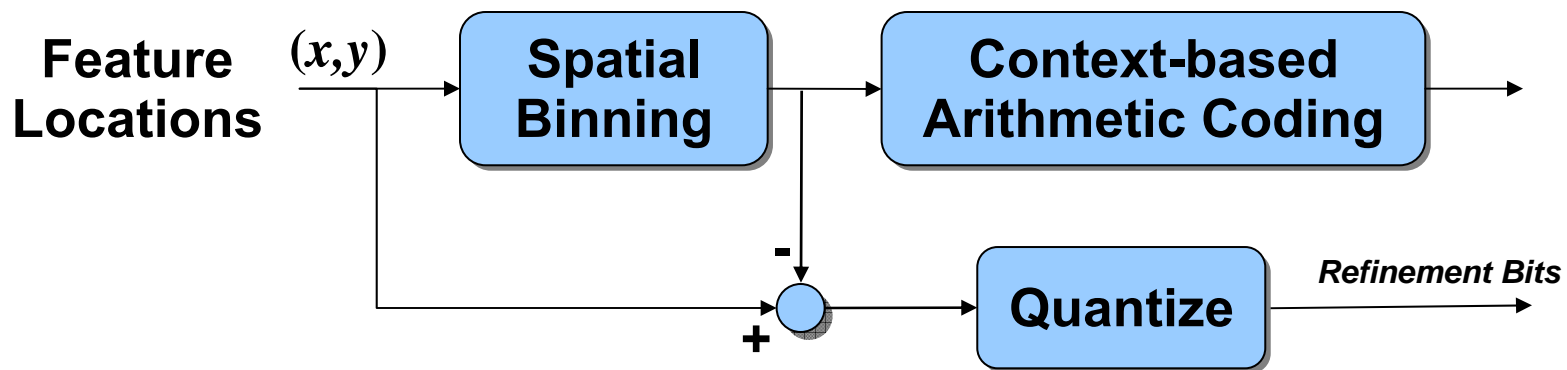
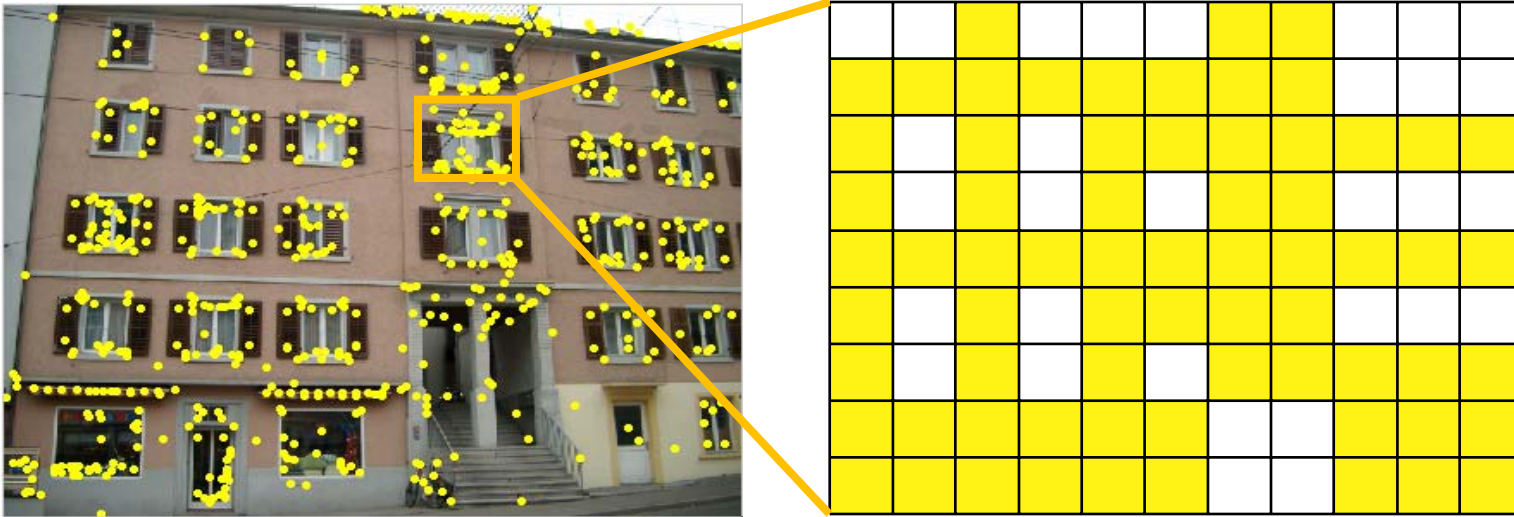


[Chandrasekhar et al., VCIP 2009]

CHoG: Compressed Histogram of Gradients

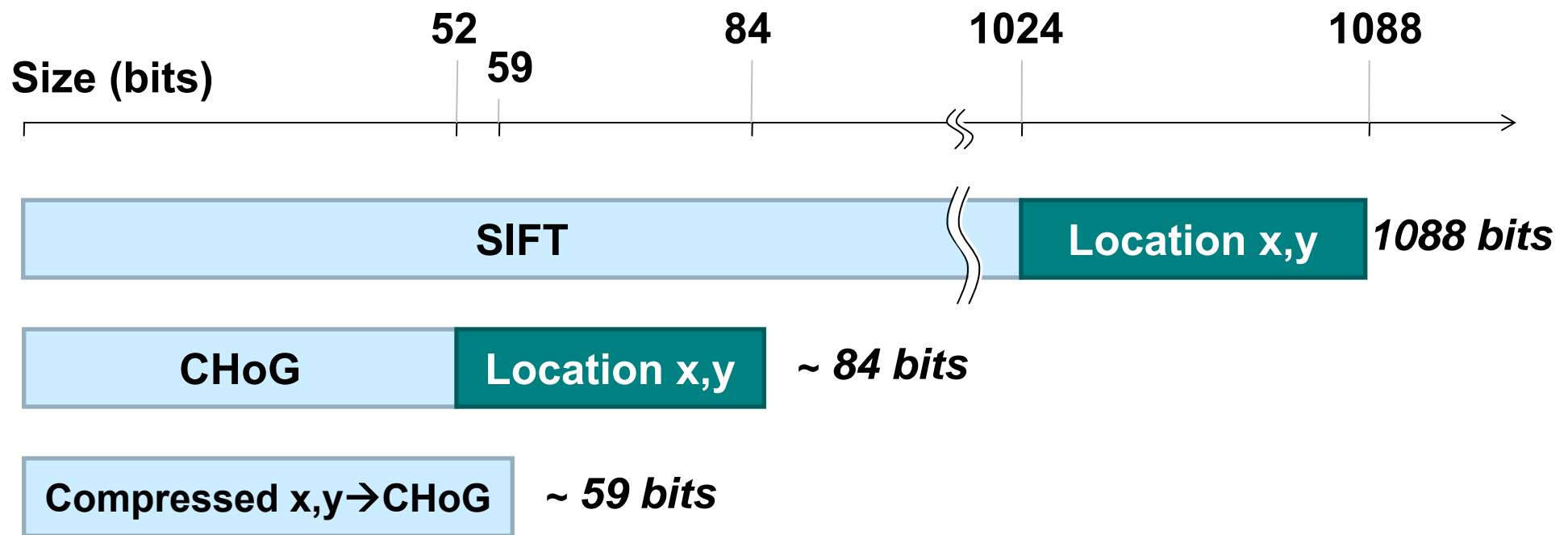


Location Histogram Coding



[Tsai et al., MobiMedia 2009]

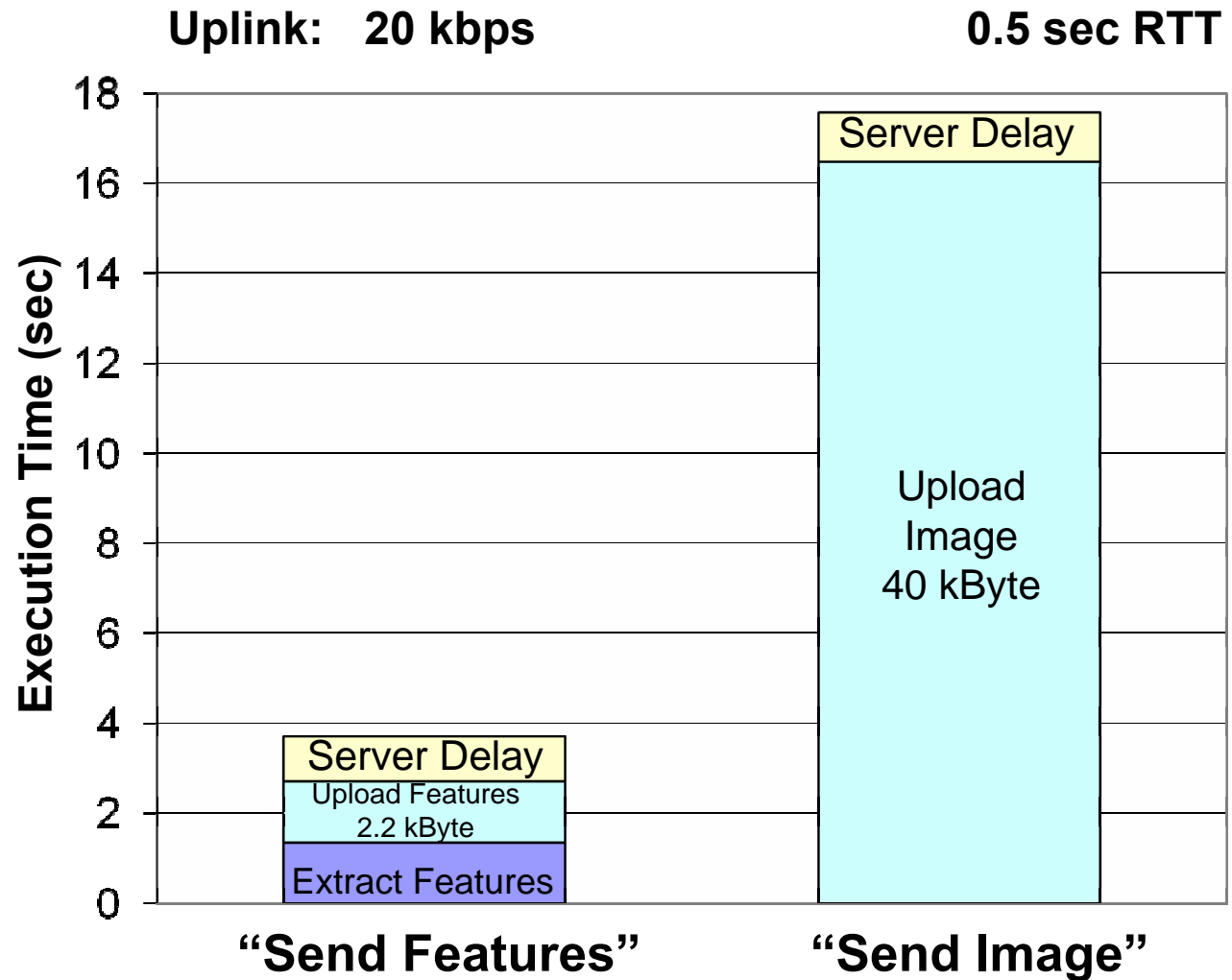
Compressed Feature Vector



[Tsai et al., MobiMedia 2009]

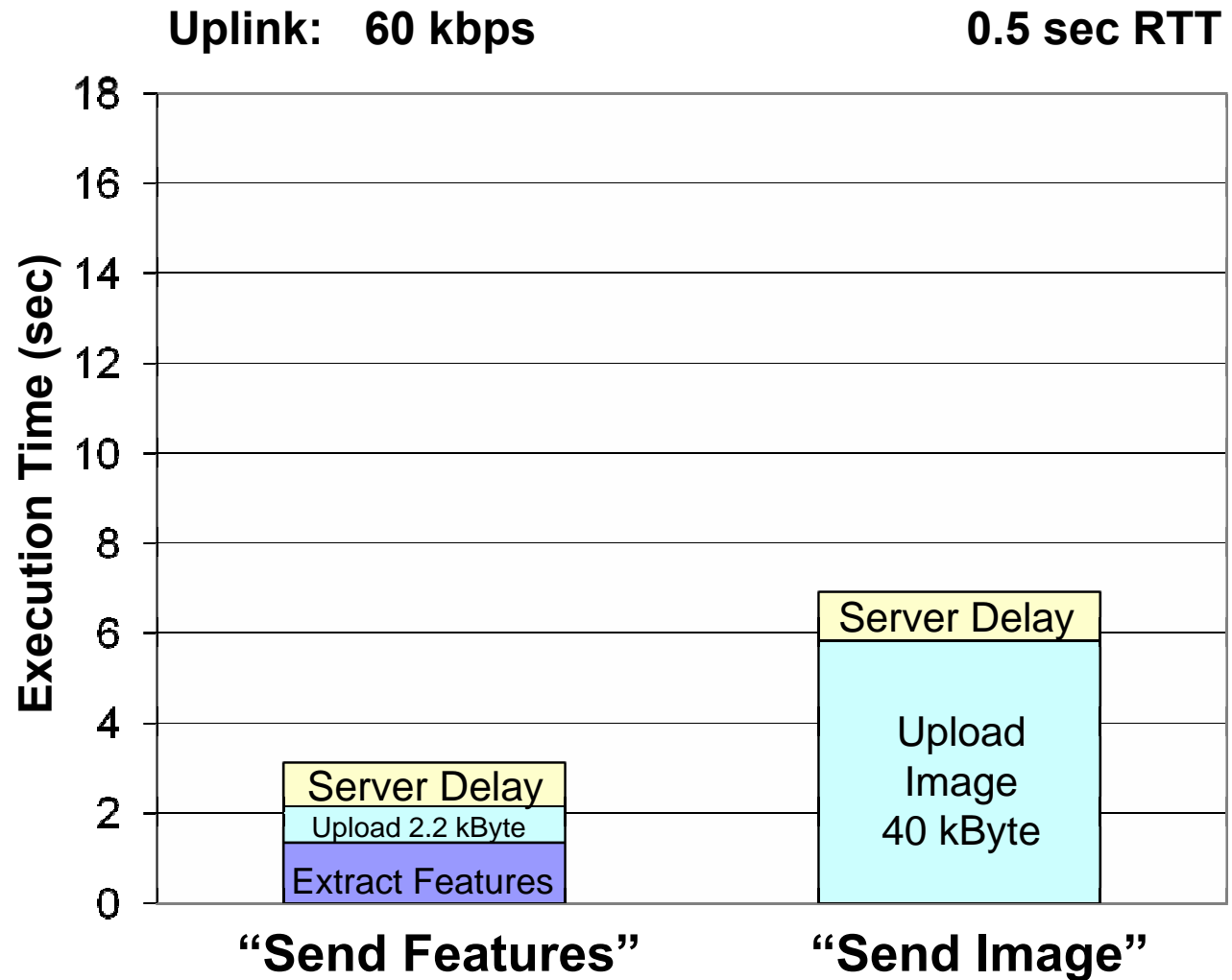
Timing Analysis

Nokia N95
330 MHz ARM
64 MB RAM



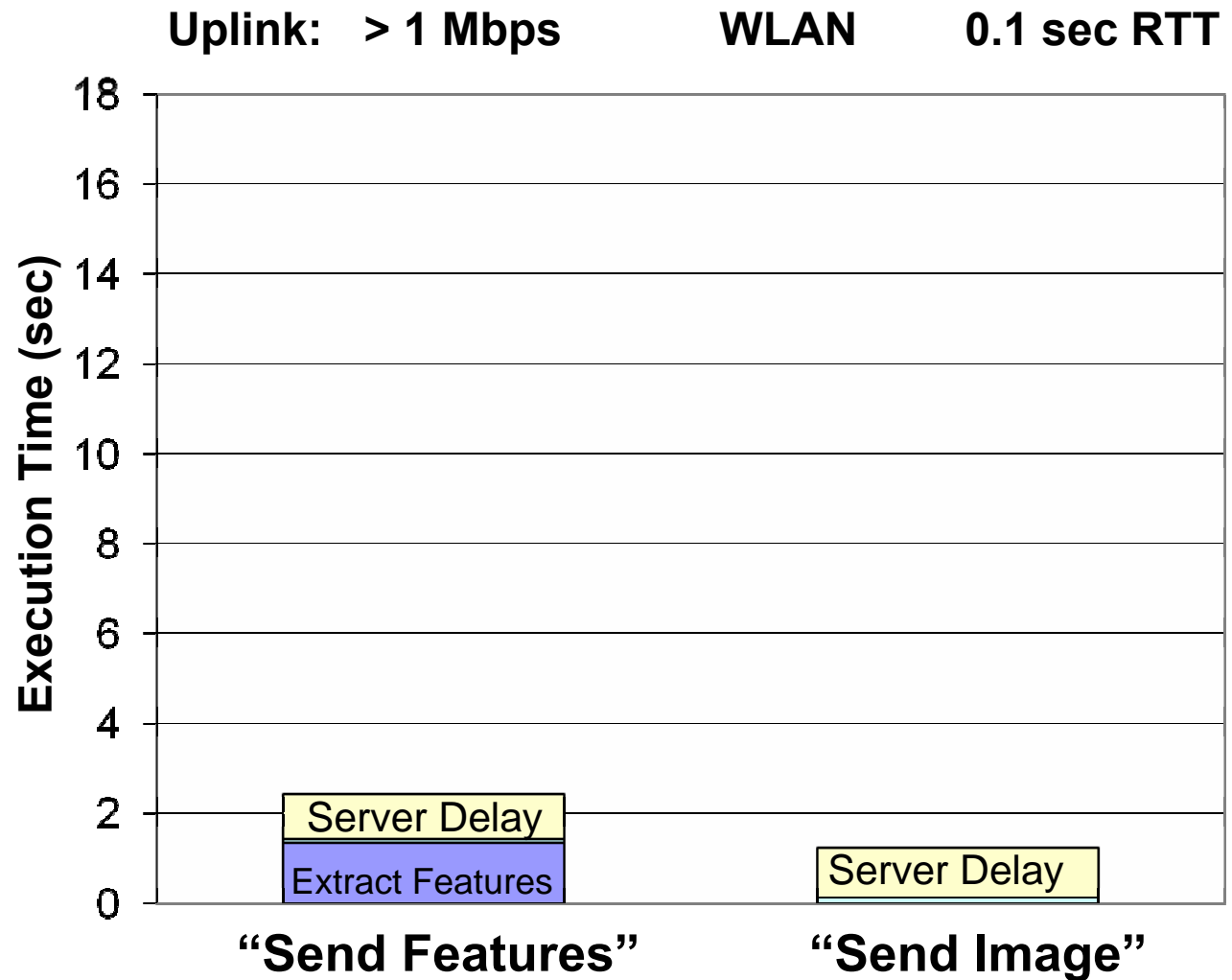
Timing Analysis

Nokia N95
330 MHz ARM
64 MB RAM

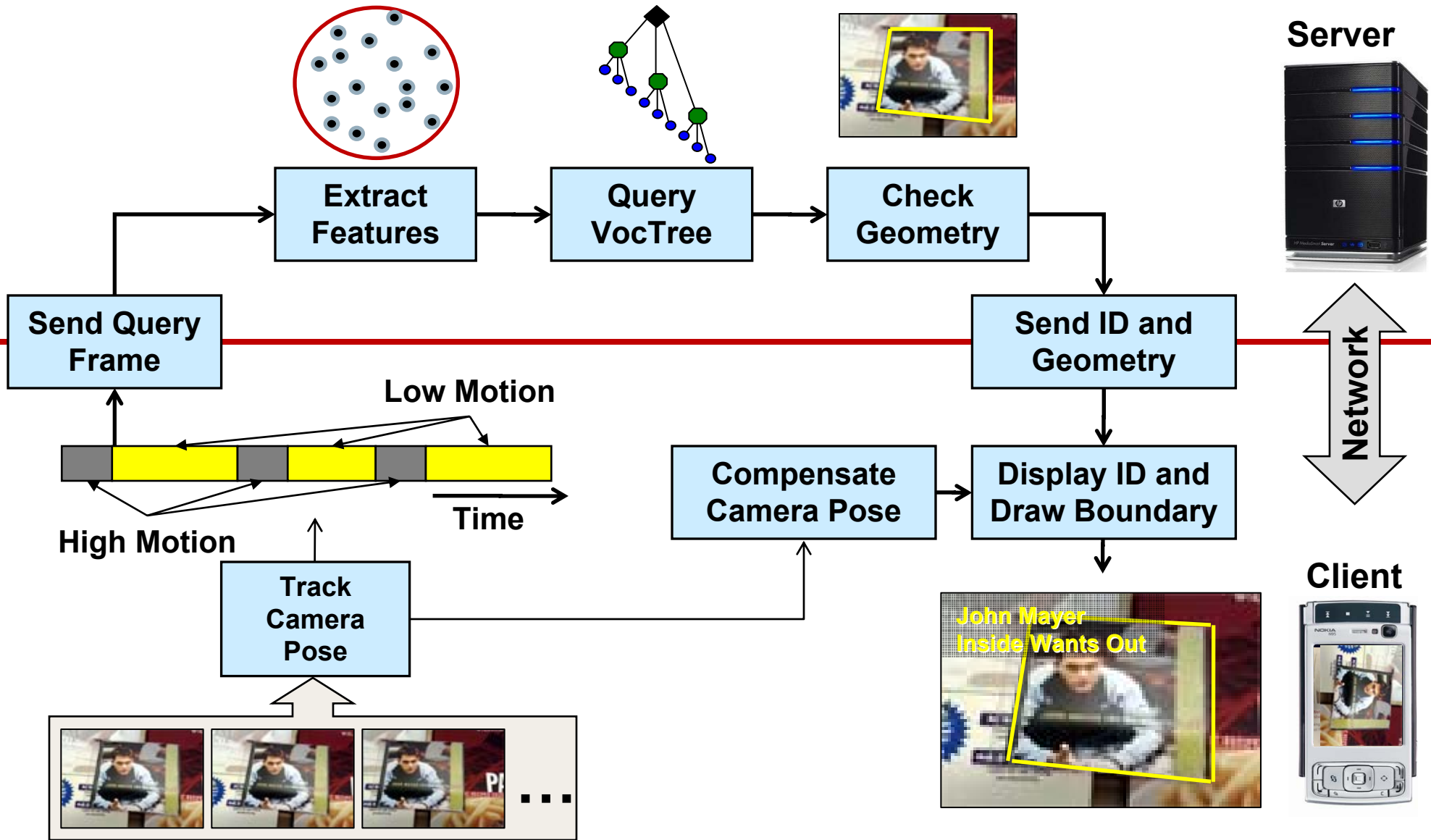


Timing Analysis

Nokia N95
330 MHz ARM
64 MB RAM



Mobile Augmented Reality



Mobile Augmented Reality: Books



Mobile Augmented Reality: CDs



Concluding Remarks

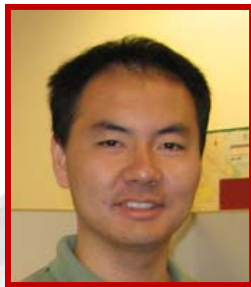
- Mobile Visual Search is ready for prime-time:
 - Bag of Words approach: > 95% recall for > 1M database
 - Vocabulary trees for fast retrieval (1 sec for > 1M database)
- Feature compression is a key problem
 - Sending a JPEG image over wireless link can be rather slow
→ send salient image features to server
 - Small database: send compressed database features to phone
 - CHoG outperforms SIFT and SURF in rate-constrained recognition
performance: ~60 bits incl. location
 - Feature extraction ~1 sec on 330 MHz smartphone
 - MPEG is considering standardization

Acknowledgments

QUALCOMM



Yuriy Reznik



David Chen



Vijay Chandrasekhar



Radek Grzeszczuk

NOKIA
Connecting People



Kari Pulli

Deutsche Telekom
Laboratories



Sam Tsai



Gabriel Takacs



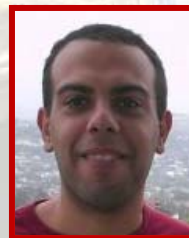
Wei-Chao Chen



Rahul Swaminathan



Jatinder Singh



Mina Makar



Thanos Bispigiannis



Natasha Gelfand



Yingen Xiong