Motivation

Where is my W-2 Form?
Video-based Tracking

Overhead video camera

Camera view of the desk
Example

40 minutes, 1024x768 @ 15 fps
System Overview

Input Video

User
System Overview

Input Video

Video Analysis

Internal Representation

Vision Engine

User
System Overview

Query Interface

Internal Representation

- Desk T
- Desk T+1

Query (where is my W-2 form?)

User

Input Video

Video Analysis
System Overview

Internal Representation

Query Interface

Query (where is my W-2 form?)

User
System Overview

Input Video

Video Analysis

Internal Representation

Vision Engine

User

T

Desk

T+1

Desk
Vision Problem
Vision Problem

... Event ...

...
Vision Problem

Event

...
Vision Problem

Event

... tut-article.pdf
... objectspaces.pdf

Desks

sanders01.pdf
lowe04sift.pdf
kidd94.pdf

Desk

Desk
Vision Problem

Scene Graph (DAG)
- tut-article.pdf
- objectspaces.pdf

Event
- sanders01.pdf
- lowe04sift.pdf
- kidd94.pdf

Desk
Assumptions

• Simplifying
  – Corresponding electronic copy exists
Assumptions

- Simplifying
  - Corresponding electronic copy exists
  - 3 event types: move/entry/exit
  - One document at a time
  - Only topmost document can move
  - No duplicate copies of same document
Assumptions

• Simplifying
  – Corresponding electronic copy exists
  – 3 event types: move/entry/exit
  – One document at a time
  – Only topmost document can move
  – No duplicate copies of same document

• Other
  – Desk need not be initially empty
Algorithm Overview

Input
Frames

Event
Detection

Document
Recognition

Scene Graph
Update
Algorithm Overview

- Input Frames
- Event Detection
- Document Recognition
- Scene Graph Update
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

“Document moved from \((x_1, y_1)\) to \((x_2, y_2)\)”

Scene Graph Update
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“A document moved from (x₁, y₁) to (x₂, y₂)”
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“A document moved from \((x_1, y_1)\) to \((x_2, y_2)\)”
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“A document moved from \((x_1,y_1)\) to \((x_2,y_2)\)”
Event Detection
Event Detection
Event Detection

Event Frames

Frame Difference

Threshold

...
Event Detection

Event Frames

... before after ...

before

after
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“A document moved from $(x_1, y_1)$ to $(x_2, y_2)$”
Event Interpretation

Move

before

after
Event Interpretation

Move

before

after

Entry
Event Interpretation

Move
before after

Entry

Exit
Event Interpretation

Move

Entry

Exit

Motion: \((x, y, \theta)\)
Event Interpretation

1. Move vs. Entry/Exit

Move

before

after

Entry

Exit
Event Interpretation

2. Entry vs. Exit

Move

Entry

Exit

before after
Event Interpretation

• Use SIFT [Lowe 99]
  – Scale Invariant Feature Transform
  – Distinctive feature descriptor
  – Reliable object recognition
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit before after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

before

after
Move vs. Entry/Exit

Motion: \((x, y, \theta)\)

before

after
Entry vs. Exit

Example 1 (entry)

before

after
Entry vs. Exit

Example 1 (entry)

before

after
Entry vs. Exit

Example 1 (entry)

before

after
Entry vs. Exit

before


after
Entry vs. Exit

Example 2 (entry)

before

after
Entry vs. Exit


before  after
Entry vs. Exit

before

after
Entry vs. Exit
Entry vs. Exit

before

after
Entry vs. Exit

before

....

after
Entry vs. Exit

Amount of change

before

……

after
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“A document moved from \((x_1,y_1)\) to \((x_2,y_2)\)”
Document Recognition

- Match against PDF image database
Document Recognition

• Match against PDF image database


• Performance
  – Can differentiate between ~100 (or more) documents
  – ~200x300 pixels per document for reliable match
Algorithm Overview

Input Frames

Event Detection

Event Interpretation

Document Recognition

Scene Graph Update

“\(A\) document moved from \((x_1, y_1)\) to \((x_2, y_2)\)”
Scene Graph Update

Motion: (x,y,θ)
Scene Graph Update

Motion: (x,y,θ)

before

after

Desk
Scene Graph Update

Motion: (x,y,θ)

before

after

Desk

Desk
Scene Graph Update

Motion: (x, y, θ)

before

after

? Desk
Scene Graph Update

Motion: (x, y, θ)
Scene Graph Update
Scene Graph Update

...
Photo Sorting Example

![Photo Sorting Example Image]
Current Directions

• Handle more realistic desktops
• Speed up processing time
• Other useful functionalities
  – Written annotation
  – Version management
  – Bookmark
  – Multi-user queries
For More Information

• Our publications

• Other related work
For More Information

• Our publications

• Other related work