Video-based Document Tracking: Unifying your Physical and Electronic Desktops. Kim, et al. UIST. 2004.

What are the core research questions addressed by the work?
- Providing a way to unify physical and electronic desktops without the need for a specialized physical infrastructure except for a video camera

What motivates the work?
- Physical paper and electronic documents both have their advantages and disadvantages. The decoupling of physical and electronic versions makes it difficult to fully exploit the advantage of both mediums. An automated system that bridges this gap is thereby needed.

How does the work understand the usage, capabilities, and limitations of paper?
- Paper's capabilities: A natural interface that people tacitly know how to interact with, preferred medium for reading, navigation and annotation, physical presence makes it easy to lay out, sort and organize multiple documents
- Paper's limitations: While memory cues from spatial layout of materials is helpful for knowledge workers, finding an exact document still difficult
- Contrasted with electronic documents
  - Advantages of electronic documents: suitable for computational operations like storage, retrieval, keyword search, sharing and version management
  - Disadvantages: Not as convenient and intuitive as direct manipulation of paper

What is the target application domain of the work?
- Specifically describes two scenarios
  - Locating documents on a desk
  - Sorting photographs

What are some proposed extensions to paper proposed by the work?
- Enabling querying of documents
- Enabling remote access
- Enabling physical manipulations to affect digital information structures

How are the proposed extensions implemented?
- Use of a fairly sophisticated vision procedure
  - Four step process: event detection, event interpretation, document recognition, scene graph update
    - SIFT highlighted as primary backbone algorithm

What findings have been obtained from either the implementation process or an evaluation of the proposed system?
- Limitations with vision based tracking:
  - Currently only enables movement of one document at a time, which does not correspond with how people interact with documents in reality
  - Does not support duplicates and poor handling of documents with a similar appearance
  - Limitations of SIFT: Recognition of documents with limited features