



# ASML

## Evaluation of Patent Image Retrieval

End-user test of PatMedia

*Has image retrieval added value over existing text based searching tools?*

# Content

- Introduction
- Test PatMedia functionality:
  - Insert *new image* (image similarity search)
  - *Figure text* search (*in description of drawings*)
  - *Patents* search → image similarity search
- Conclusions

# Introduction

- PatMedia:
  - Patent image retrieval engine
  - Developed by the Informatics and Telematics Institute in Greece
  - Presented at the IRFS 2008 by Stefanos Vrochidis
- Test data:
  - Exposure apparatus for Lithography
  - IPC G03F007/20
  - 15.000 patents, 322360 images (IRF MAREC database)
  - Uploading and image segmentation (>2 weeks)

# PatMedia interface

Informatics & Telematics Institute

## PatMedia

FIGURES: 322360 [1-24]

Insert New Image

Upload a new image and search...

Figure Text Search

Insert a keyword (e.g. laser) to search the figure descriptions...

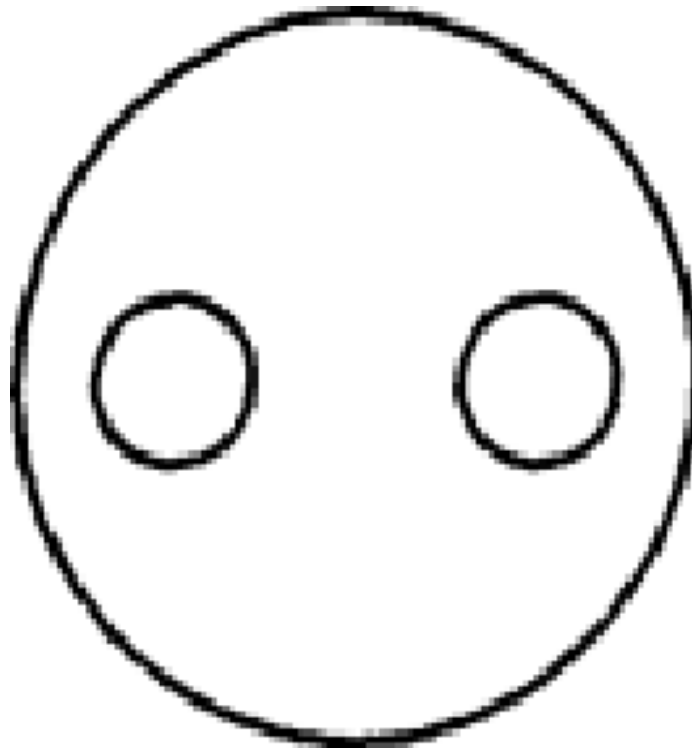
Patents Search

Give the publication number of the patent to retrieve patent images...

Text Filtering

Insert a keyword to filter the results, e.g.: disc, laser...

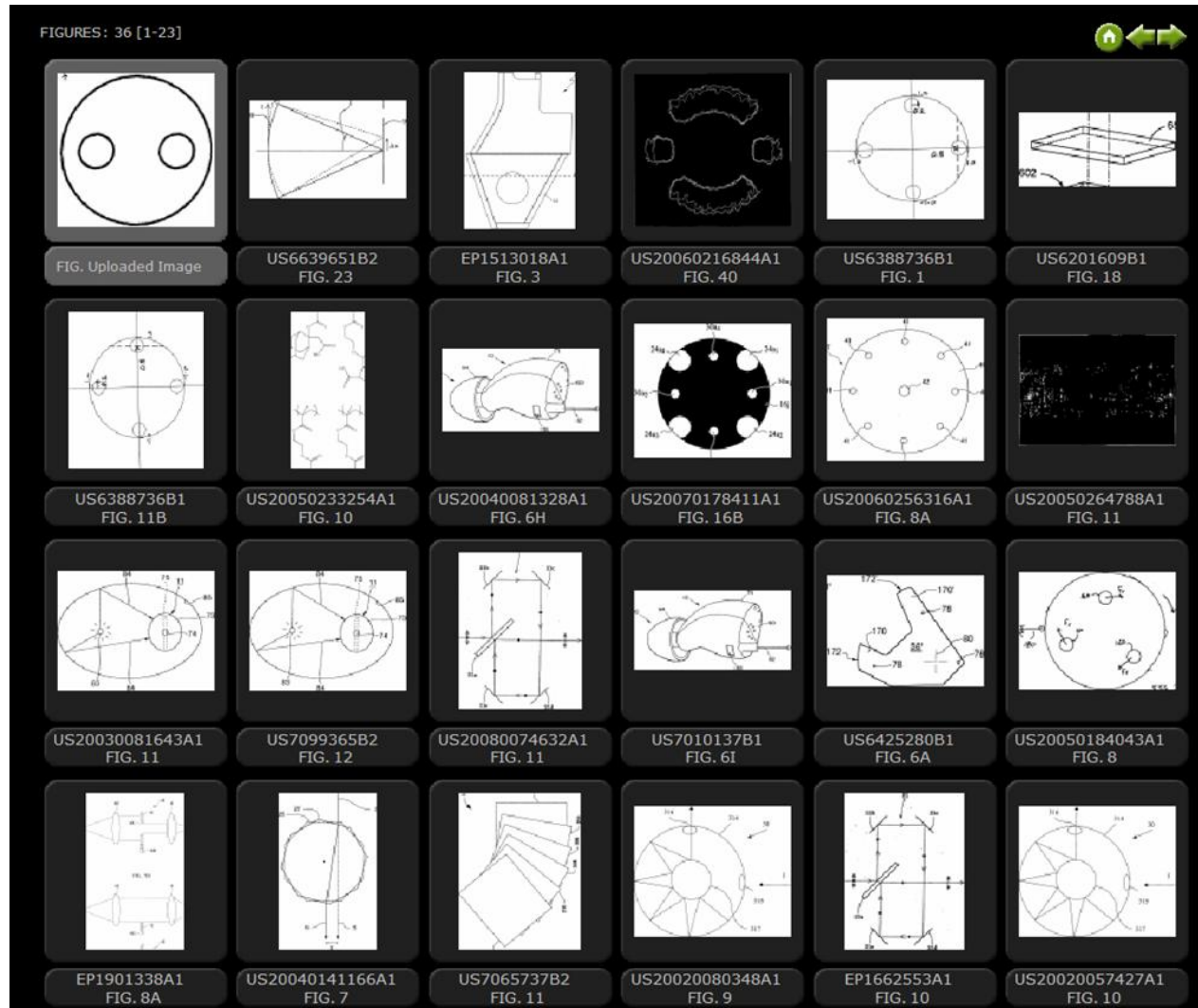
# Insert new image (image similarity search)



# Results “Insert New Image” search (1)

Visual  
Relevant:  
(1),4,5  
7,10,11  
13,14,18  
22,24

Technical  
Relevant:  
(1),4,5  
7,9

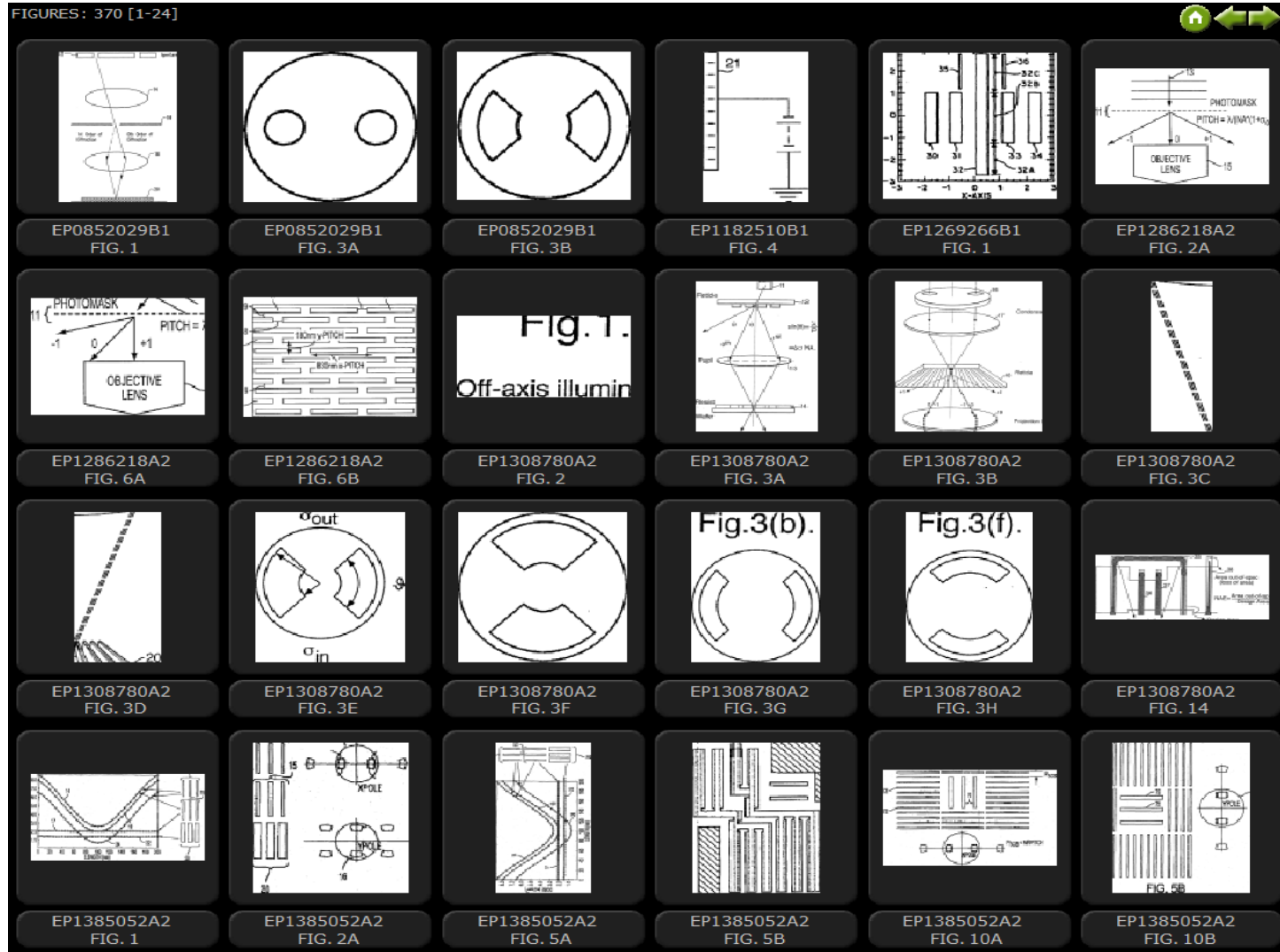


# Figure Text Search (in figure descriptions)

**Figure Text Search**

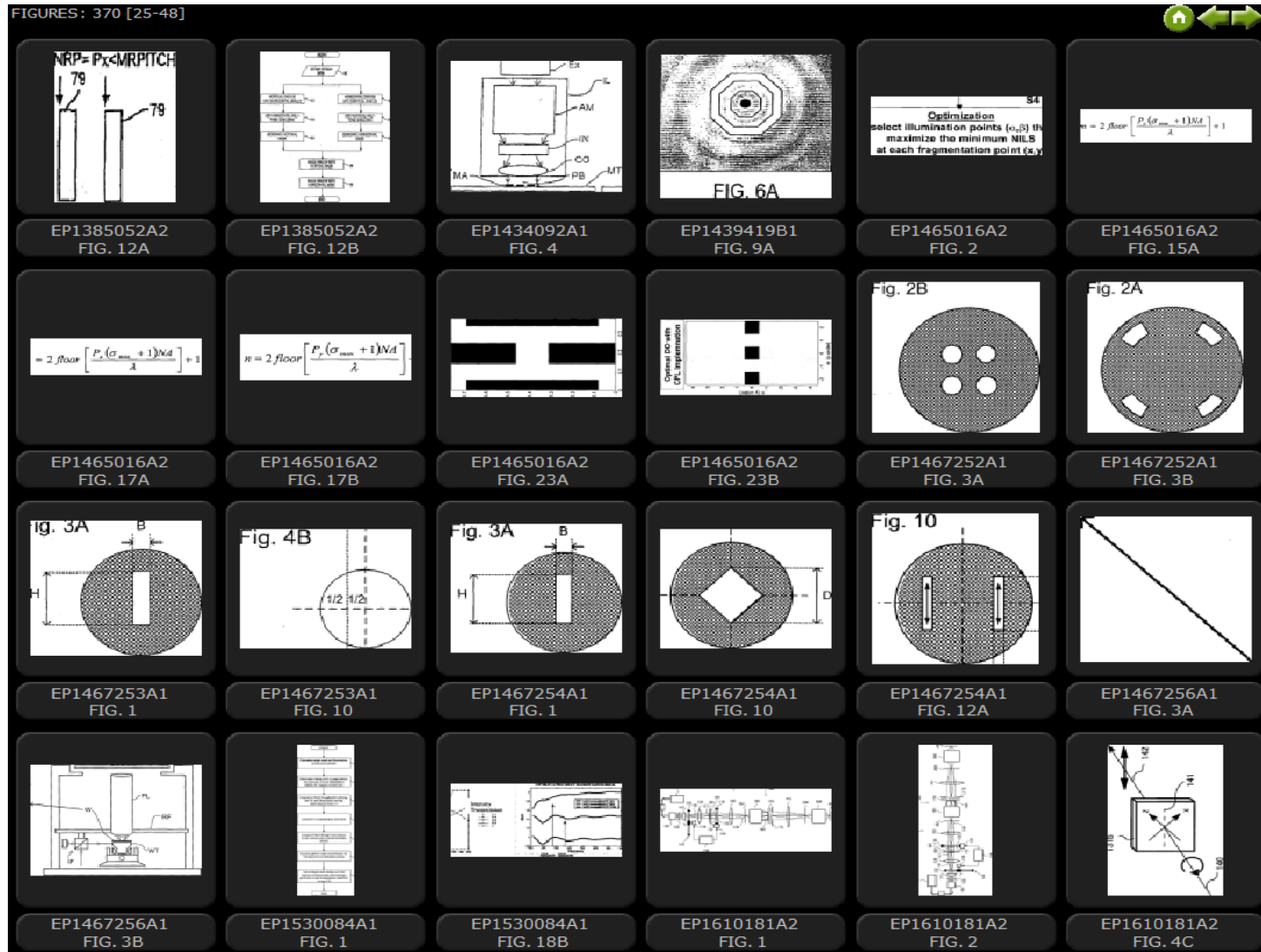
Insert a keyword (e.g. laser) to search the figure descriptions...

# Results “Dipole” Figure Text Search (1)

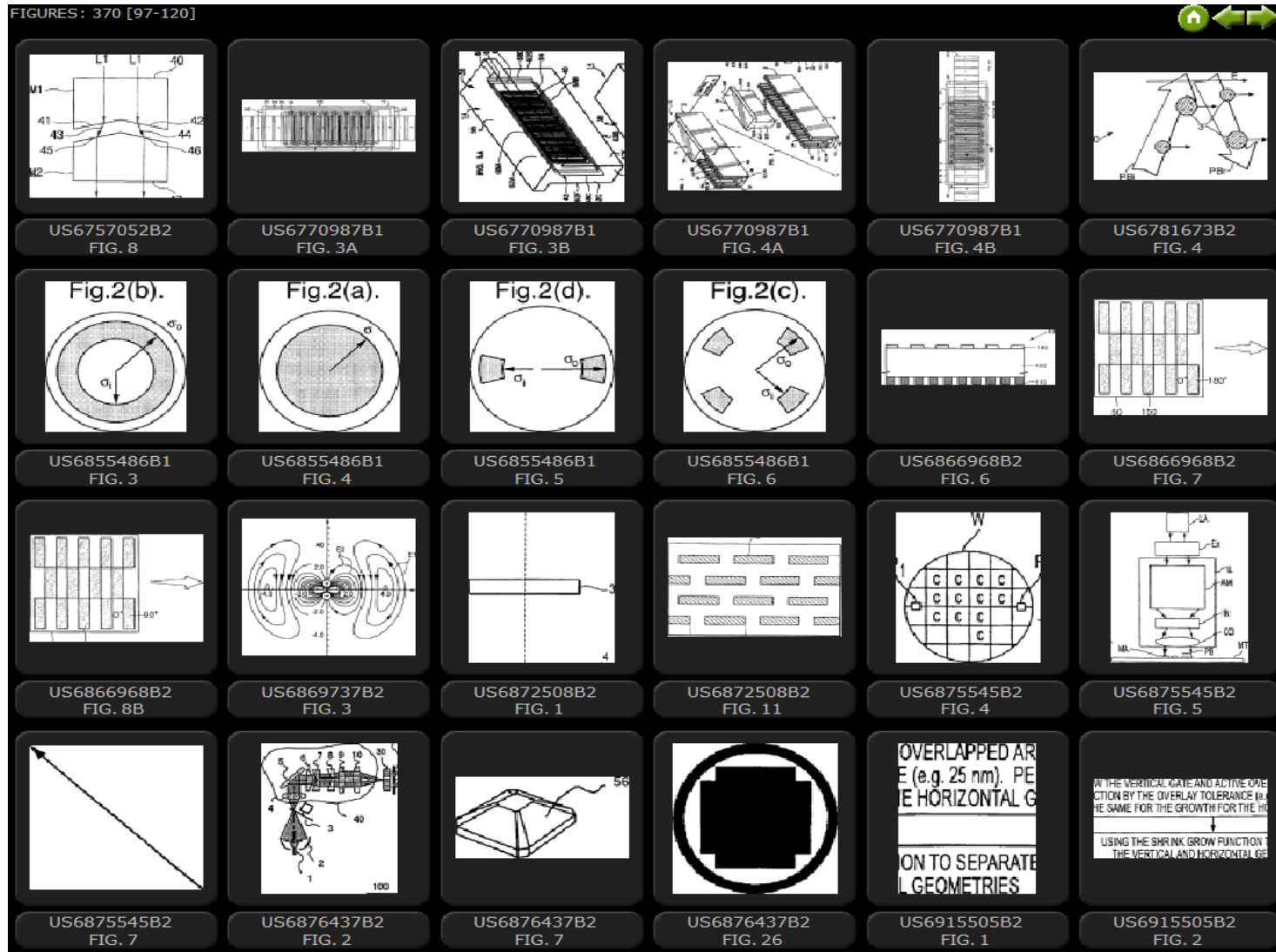




# Results “Dipole” Figure Text Search (2)



# Results “Dipole” Figure Text Search (3)



# Selection from 370 hits on “Dipole” text search

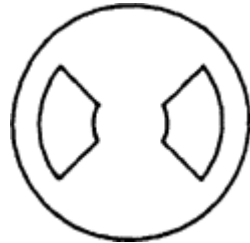


Fig.3(b).

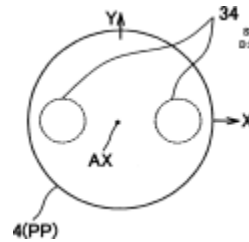


Fig.3(g).

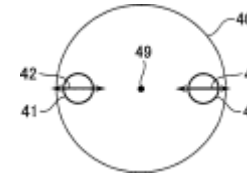


Fig.3(f).

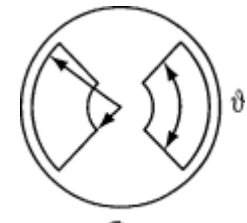
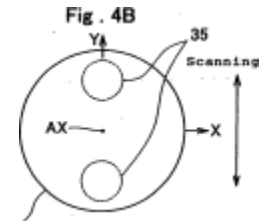


FIG. 8

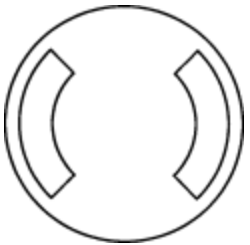
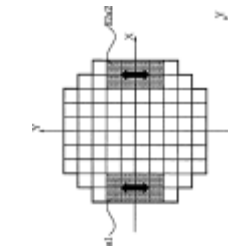
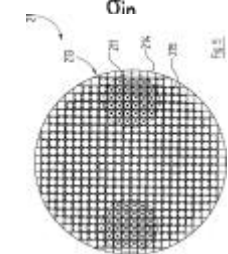
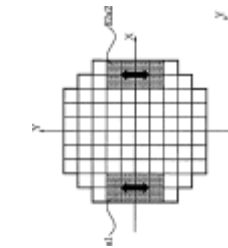
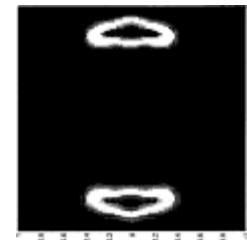
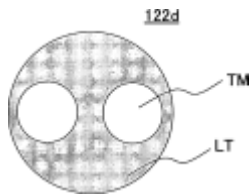
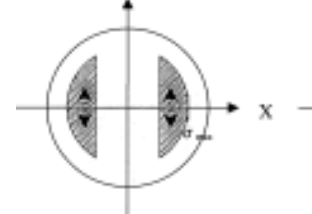
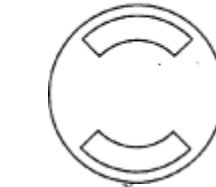
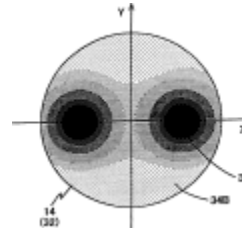
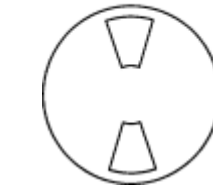
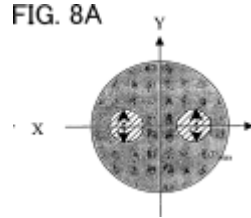


FIG. 8A

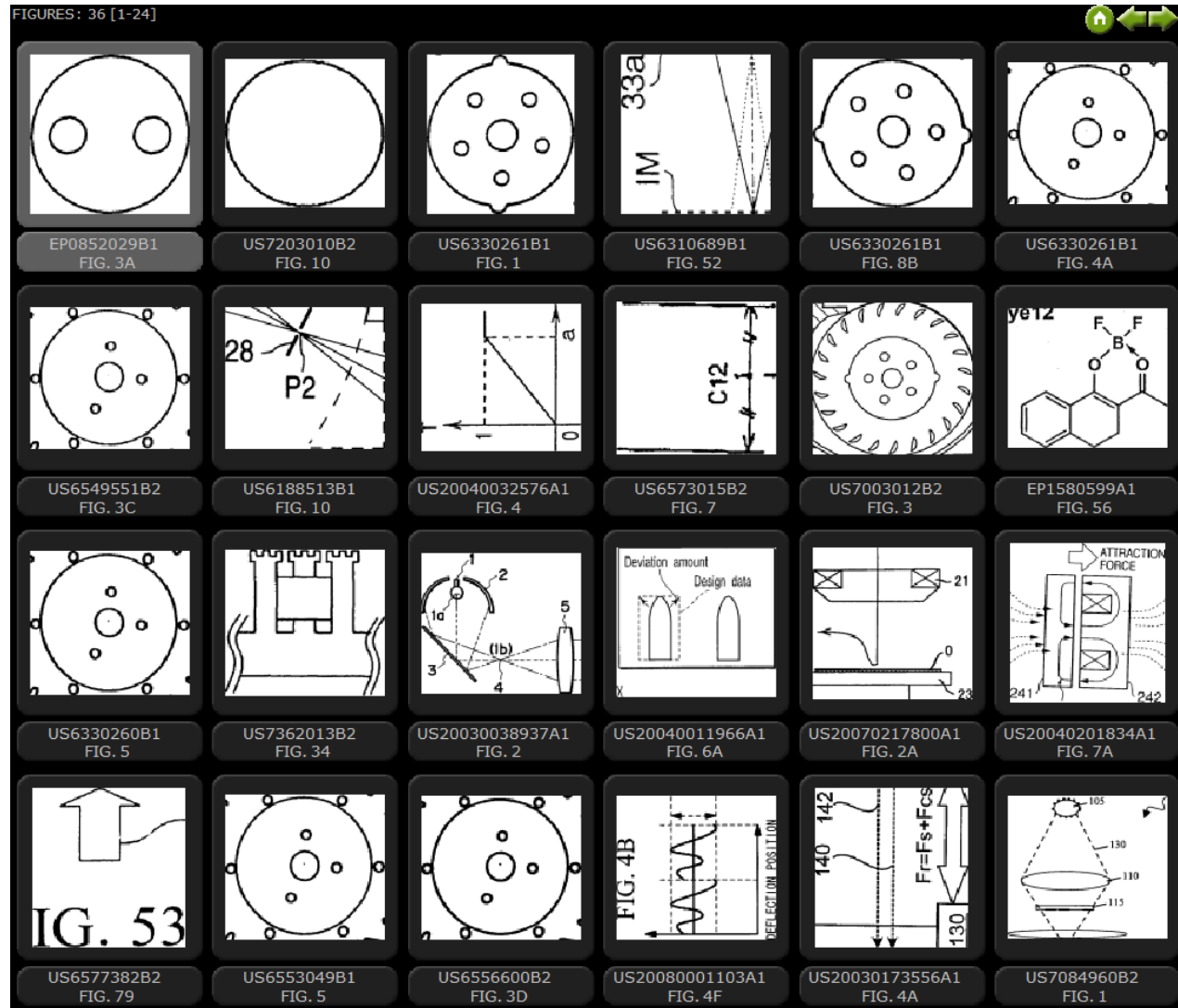


ASML

# Results Image similarity search (1) (existing drawing)

Visual  
Relevant:  
(1),2,3,5,6  
7,11  
13  
20,21

Technical  
Relevant:  
(1)



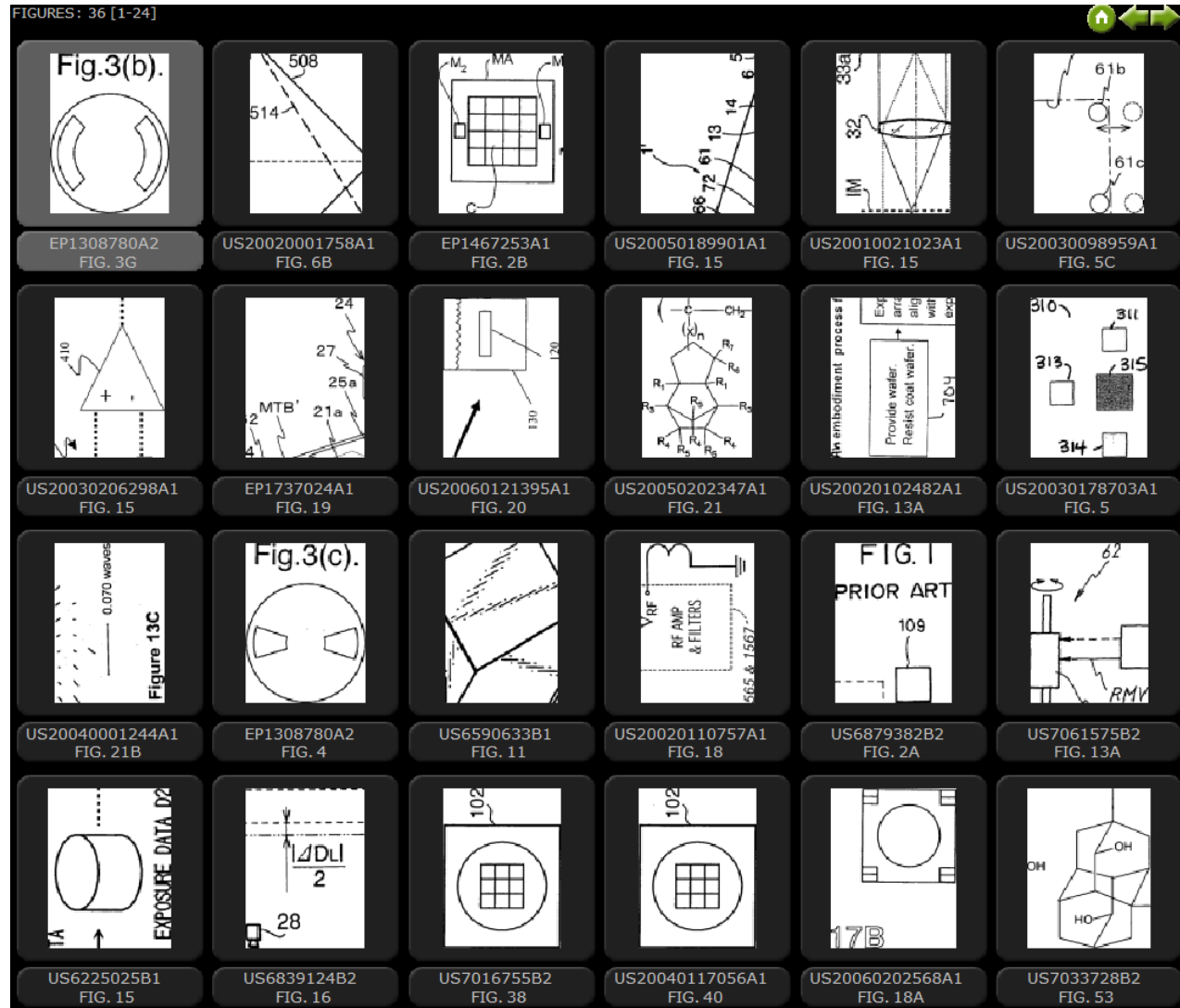
# Results Image similarity search (2) (existing drawing)

Visual  
Relevant:  
(1)

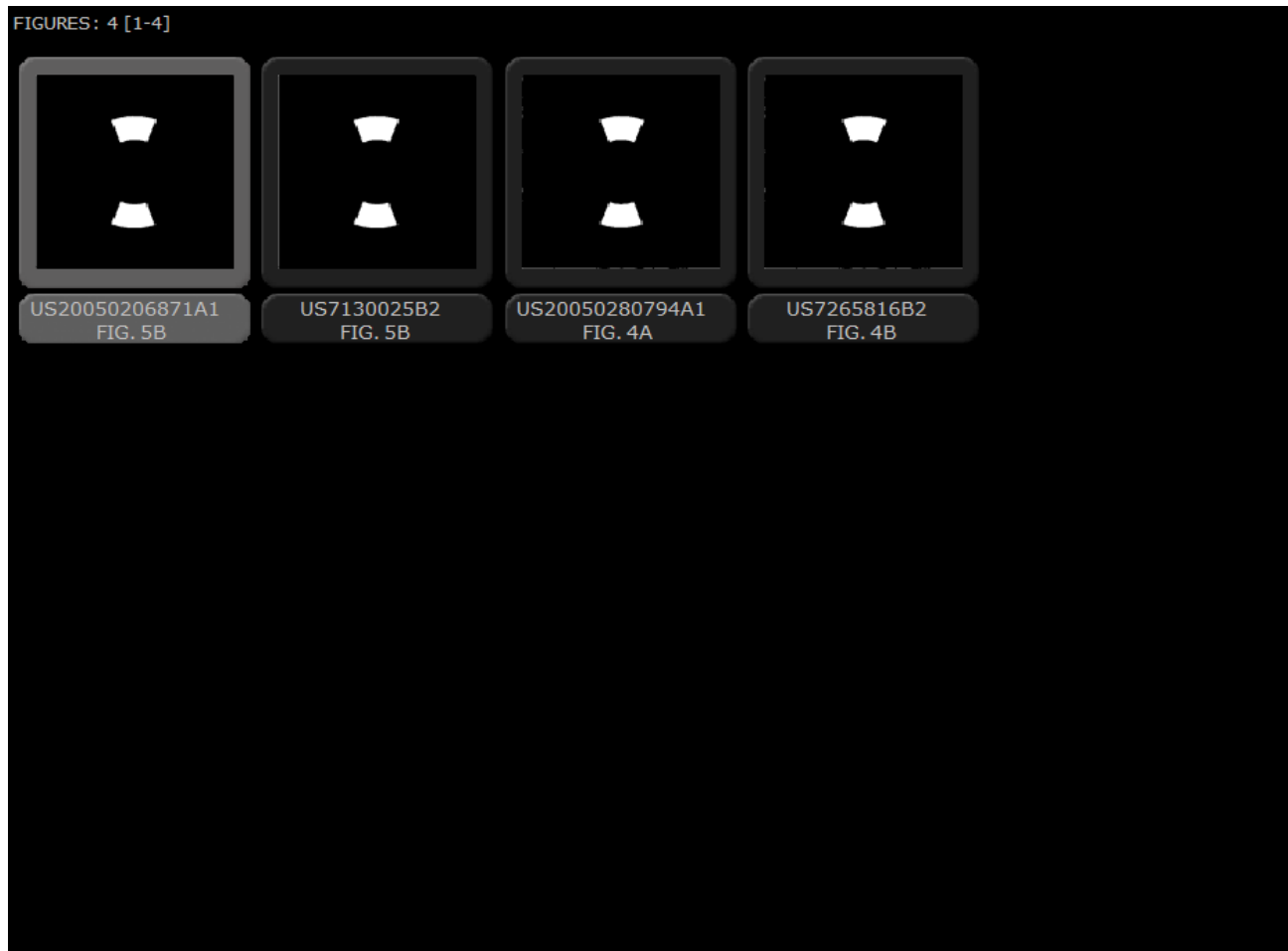
14  
21,22

Technical  
Relevant:  
(1)

14



## Results Image similarity search (3) (existing drawing)



All relevant (related family members)

# Patents Search → image similarity search

- Standard ASML image EP 1262836 A1 (tool check)
- US 6917412 B2:
  - Validity, Value
    - US-only, stand-alone image?
  - How does a drawing (technology) evolve over time?  
(sorting by publication date)



5

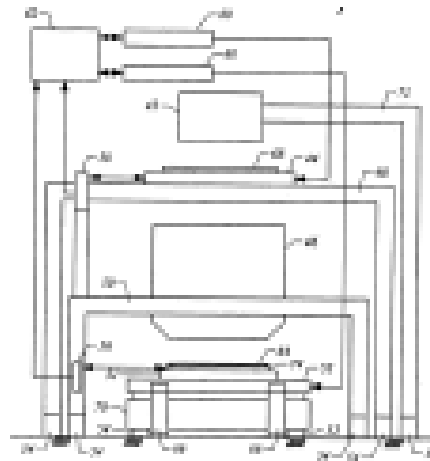
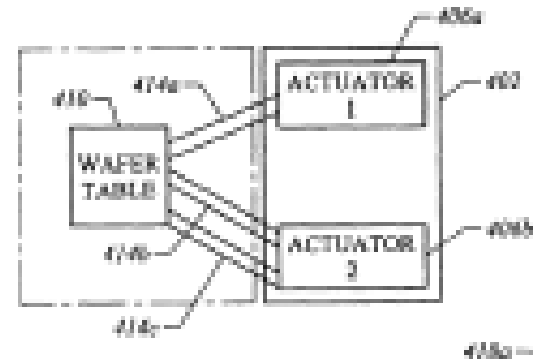
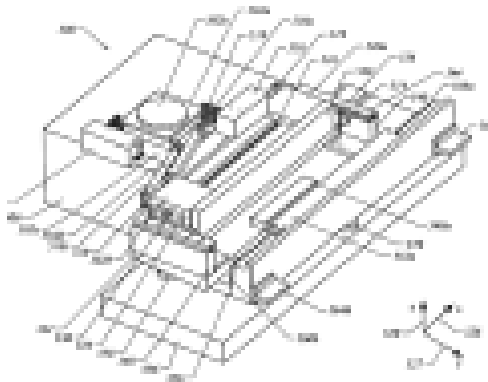




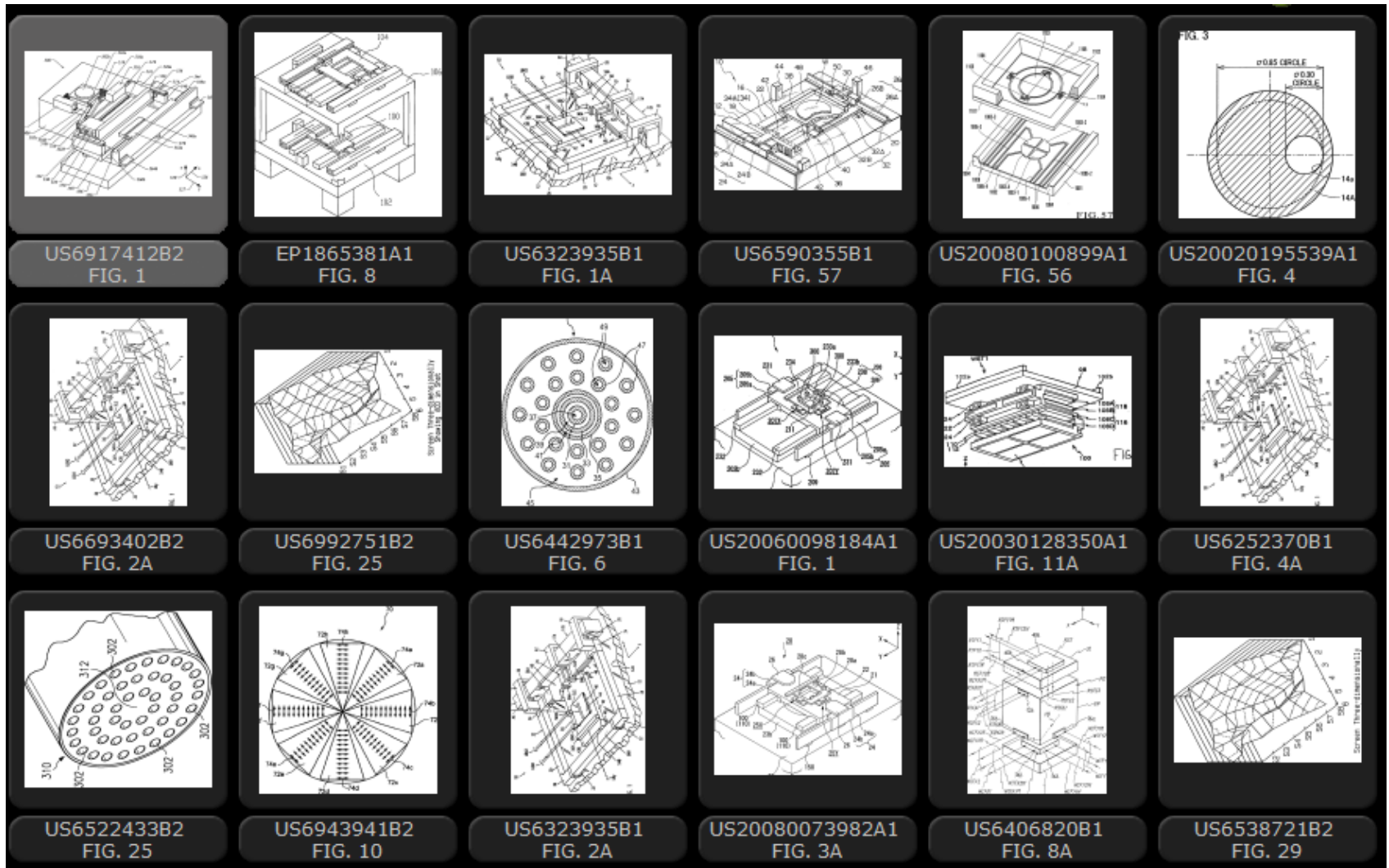
# Results for standard ASML image



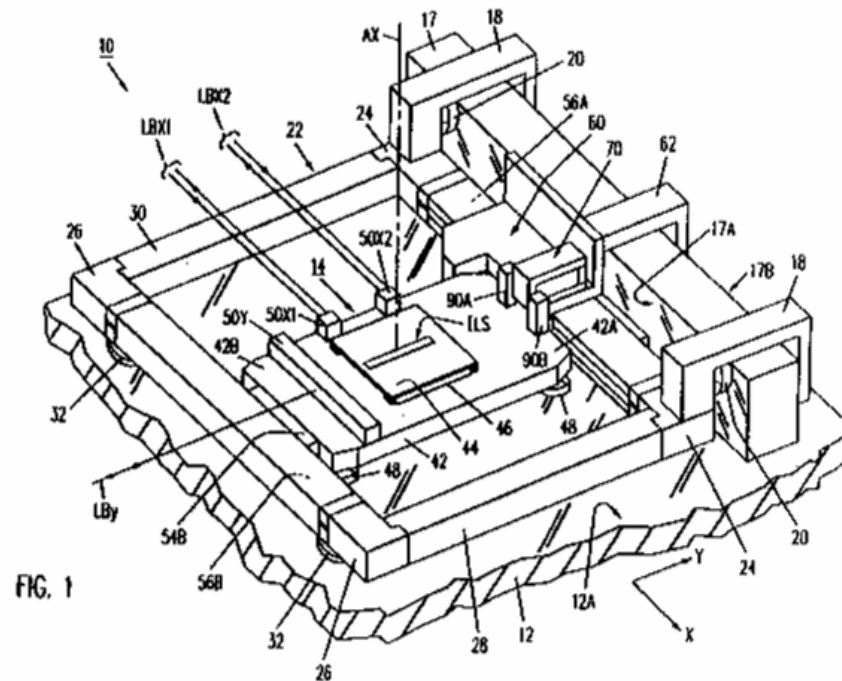
# Example figures US 6,917,412 B2



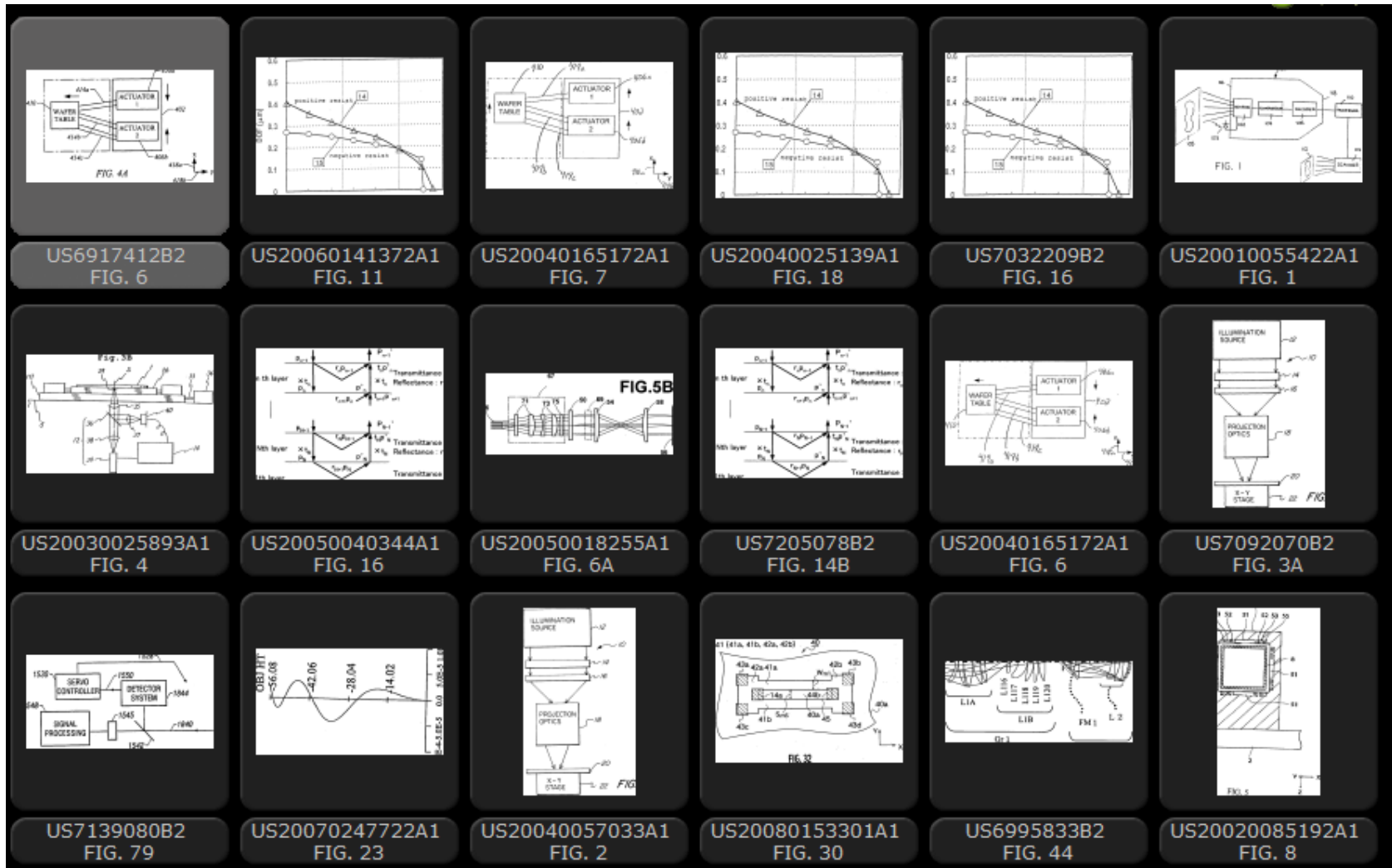
# Results figure similarity search US 6,917,412 B2 (1)



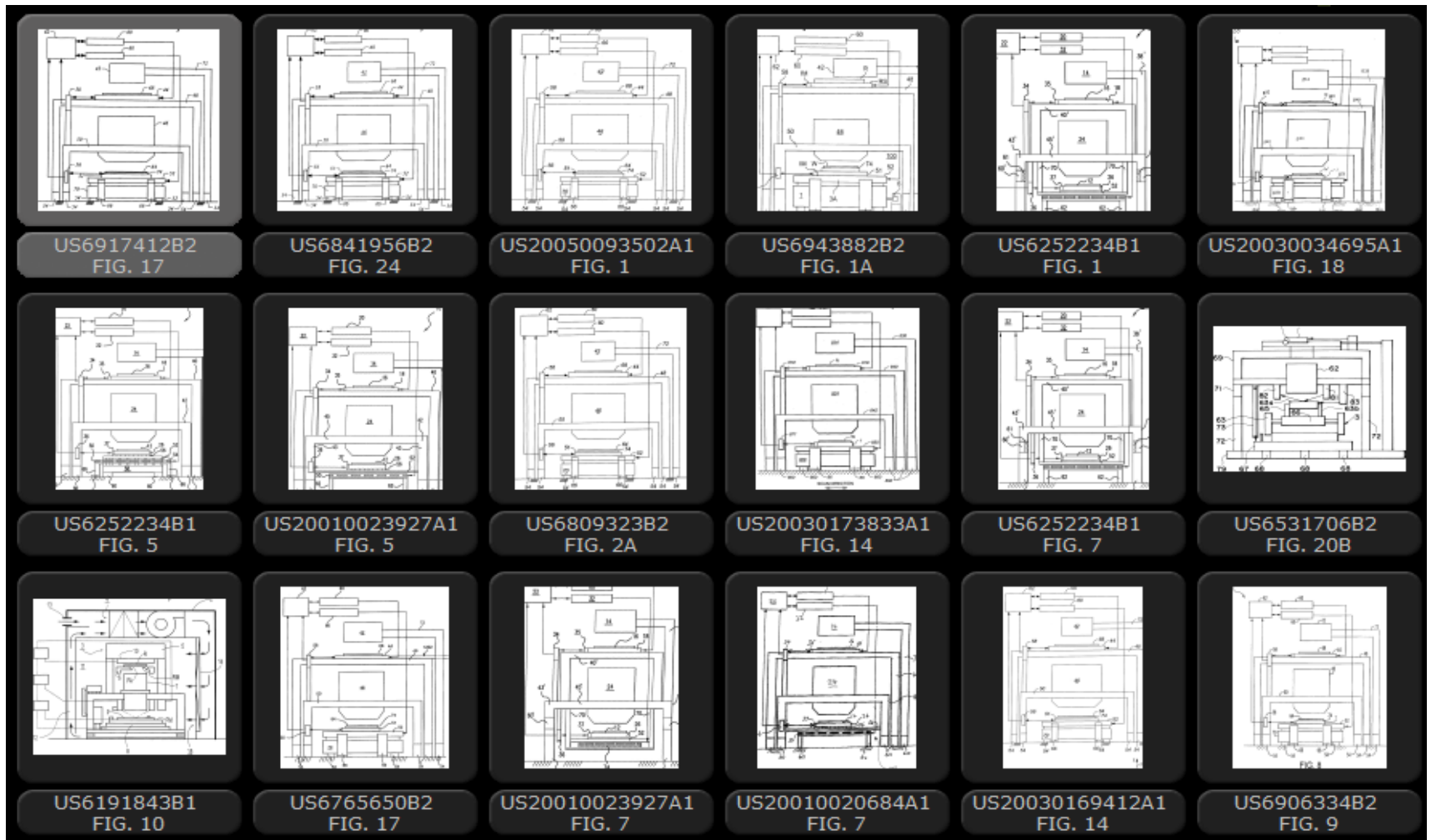
# Cited reference!



# Results figure similarity search US 6,917,412 B2 (2)



# Results figure similarity search US 6,917,412 B2 (3)



# Conclusions (tested functionality)

- Insert *new Image (similarity search)*:

Added value: Novelty, patentability searches (upload invention disclosure figure/drawing)

- Consider “synonyms” (rotated, inverted, color, frames)
- Remove pointers, text, numbers, etc.
- Needs improvement to get higher precision (i.e. less noise)

- *Figure text search (description of drawings)*:

Added value: All search types (“Google Image” search)

- Yields relevant images of the searched text instead of the patent scope!
- Relatively easy to integrate in existing text based tools: Requires segmented figures, with linked searchable figure description)
- Boolean logic: Truncation, AND/OR

- *Patents (number) search* → image similarity search:

Added value:

- Validity/Invalidity searching (quick check, limited effort)
- Patent valuation (first/only/many time used?)
- Competitive intelligence research (evolution of drawings)

# Conclusions (General)

- Language independent
- Fast culling (370 hits within minutes)
- Result presentation could be improved with:
  - Sorting (date, relevance, PN,CC, PA, etc.)
  - Decuplication (one hit per patent family)
  - Figure selection with export functionality
  - .....
- Training/relevance feed back with selected images