DEMONSTRATIVES OF PETITIONERS CANON INC., CANON U.S.A., INC. AND AXIS COMMUNICATIONS AB

Oral Hearing

IPR2019-00311 and IPR2019-00314

April 8, 2020

U.S. Patent No. 7,932,923

Overview

’923 Patent Claims Are Unpatentable

Overview of ’923 Patent and Claimed Invention

IPR2019-00311: Anticipated by Kellogg or Obvious over Kellogg and Brill

IPR2019-00314: Obvious over Dimitrova and Brill

Patent Owner admits it did not invent new object detection or computer vision techniques

Although the video surveillance system of the invention draws on well-known computer vision techniques from the public domain,

IPR2019-00311, Ex. 1001, 5:6-8

The following references describe moving target detection: ...

The following references describe detecting and tracking humans: ...

The following references describe blob analysis: ...

The following references describe blob analysis for trucks, cars, and people: ...

The following reference describes analyzing a single-person blob and its contours: ...

The following reference describes internal motion of blobs, including any motion-based segmentation: ...

IPR2019-00311, Ex. 1001, 1:27-2:18

As an option, blocks **51-54** can be replaced with any detection and tracking scheme, as is known to those of ordinary skill.

IPR2019-00311, Ex. 1001, 10:27-29

Classification can be performed by a number of techniques, and examples of such techniques include using a neural network classifier {14} and using a linear discriminatant classifier {14}. Examples of classification are the same as those discussed for block **23**.

IPR2019-00311, Ex. 1001, 10:44-48

In block **57**, video primitives are identified using the information from blocks **51-56** and additional processing as necessary. Examples of video primitives identified are the same as those discussed for block **23**. As an example, for size, the system can use information obtained from calibration in block **22** as a video primitive. From calibration, the system has sufficient information to determine the approximate size of an object. As another example, the system can use velocity as measured from block **54** as a video primitive.

IPR2019-00311, Ex. 1001, 10:49-57

Disclosed Invention – Detecting video primitives and applying an event discriminator to the video primitives

(57) **ABSTRACT**

A video surveillance system is set up, calibrated, tasked, and operated. The system extracts video primitives and extracts event occurrences from the video primitives using event discriminators. The system can undertake a response, such as an alarm, based on extracted event occurrences.

IPR2019-00311, Ex. 1001, 1

An operator is provided with maximum flexibility in configuring the system by using event discriminators. Event discriminators are identified with one or more objects (whose descriptions are based on video primitives), along with one or more optional spatial attributes, and/or one or more optional temporal attributes. For example, an operator can define an event discriminator (called a "loitering" event in this example) as a "person" object in the "automatic teller machine" space for "longer than 15 minutes" and "between 10:00 p.m. and 6:00 a.m."

IPR2019-00311, Ex. 1001, 4:63-5:5

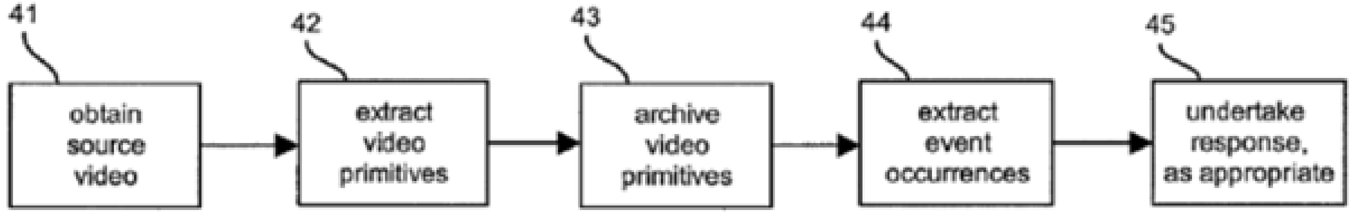


FIG. 4

Ex. 1001, Fig. 4

Disclosed Invention – Video primitives/attributes include motions and activities.

An event discriminator is described in terms of video primitives. A video primitive refers to an observable attribute of an object viewed in a video feed. Examples of video primitives include the following: a classification; a size; a shape; a color; a texture; a position; a velocity; a speed; an internal motion; a motion; a salient motion; a feature of a salient motion; a scene change; a feature of a scene change; and a pre-defined model.

IPR2019-00311, Ex. 1001, 7:5-12

A motion refers to any motion that can be automatically detected. Examples of a motion include: appearance of an object; disappearance of an object; a vertical movement of an object; a horizontal movement of an object; and a periodic movement of an object.

A salient motion refers to any motion that can be automatically detected and can be tracked for some period of time. Such a moving object exhibits apparently purposeful motion. Examples of a salient motion include: moving from one place to another; and moving to interact with another object.

A feature of a salient motion refers to a property of a salient motion. Examples of a feature of a salient motion include: a trajectory; a length of a trajectory in image space; an approxi-

IPR2019-00311, Ex. 1001, 7:37-49

Representative Claim 1

1. A method comprising:

[1.1] detecting an object in a video from a single camera;

[1.2] detecting a plurality of attributes of the object by analyzing the video from said single camera,
the plurality of attributes including at least one of a physical attribute and a temporal attribute, each attribute representing a characteristic of the detected object;

[1.3] selecting a new user rule after detecting the plurality of attributes; and

[1.4] after detecting the plurality of attributes and after selecting the new user rule, identifying an event of the object that is not one of the detected attributes of the object by applying the new user rule to the plurality of detected attributes,

[1.5] wherein the applying the new user rule to the plurality of detected attributes comprises applying the new user rule to only the plurality of detected attributes;

[1.6] wherein the plurality of attributes that are detected are independent of which event is identified,

[1.7] wherein the step of identifying the event of the object identifies the event without reprocessing the video, and

[1.8] wherein the event of the object refers to the object engaged in an activity.

Issues Raised By Avigilon

- Attributes vs. Events (Dimitrova)
- New User Rule Requires A Response (Dimitrova)
- Applying The User Rule To Attributes
- Applying The New User Rule To Only The Attributes (Dimitrova)
- Single Camera
- Subset (Dimitrova)
- Video Device
- Obvious To Combine (Dimitrova)
- Publication Of Prior Art (Dimitrova)

The PTAB Need Not Revisit Previously Decided Issues

Collateral Estoppel Applies To Issues When:

(1) the issue is the same as the issue in the prior action;

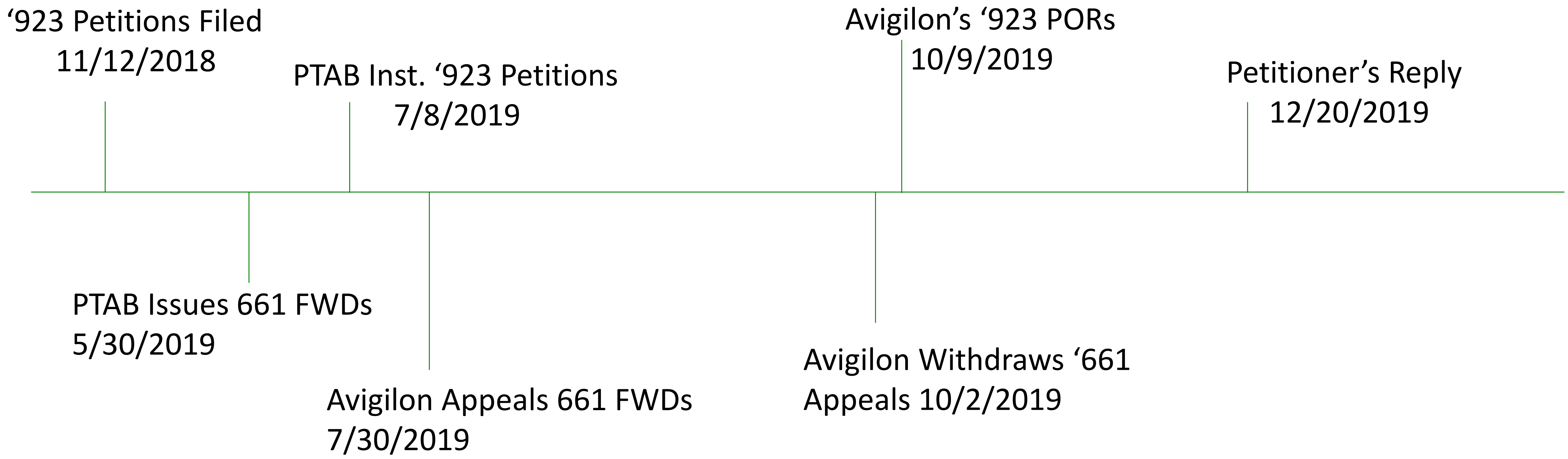
(2) the issue was actually litigated in the first action;

(3) there was a final judgement in the first action that necessarily required determination of the identical issue; and

(4) the prior action featured full representation of the estopped party.

Mobile Tech, IPR2018-00481, Final Written Decision, Paper 29 at 10.

Collateral Estoppel - Timing



Collateral Estoppel – the PTAB has already decided

- A “user rule” does not require a response
- “Independent” (Arg. 2) - “the detection of attributes is independent from, i.e., not affected by, the user rule that tasks the system.”
- Kellogg teaches the independence elements and is not like Courtney
- A POSA would be motivated to combine Kellogg and Brill
- Dimitrova teaches the independence elements and is not like Courtney
- A POSA would be motivated to combine Dimitrova and Brill

ATTRIBUTES VS. EVENTS

ATTRIBUTES VS. EVENTS

- Avigilon argues that the prior art is distinguishable because it detects events instead of attributes, which is the faulty basis for several of its arguments, see:

  - Kellogg does not disclose “the plurality of attributes that are detected are independent of which event is identified” POR at 31
  - Kellogg does not disclose “identifying…applying the new user rule to only the plurality of detected attributes” POR at 23
  - Kellogg does not disclose “selecting a new user rule after detecting the plurality of attributes” POR at 22
  - Kellogg does not disclose “wherein the memory is configured to store at least some of the plurality of attributes for at least two months” POR at 38

Kellogg stores object attributes and allows a user to define and search for *ad hoc* events

Visual memory could act as the interface between inputs and applications in a computer vision system. For example, computer vision algorithms for a security system could analyze data provided by various cameras and store information in the visual memory. Applications could then retrieve this data to track objects, watch for suspicious events, and respond to user queries. The visual memory would coordinate the information from its inputs and eliminate the need for full connectivity between inputs and applications.

IPR2019-00311 Ex. 1003 (Kellogg) at 10

KELLOGG's approach event demonstrates *ad hoc* events

```
Select p from Person
  during %times-1
  where p in
     (Select q from Person
        where q centroid within 3 of %spatiotemporal-spec
        during %times-2)
```

IPR2019-00311 Ex. 1003 (Kellogg) at 63

The PTAB has already decided that *Kellogg* discloses object and attribute detection that is independent from event determination

To the extent Patent Owner argues that a method of "detecting objects or attributes" is a "machine vision system," we find Kellogg discloses this limitation. PO Resp. 35. As discussed above, we agree with Petitioner that Kellogg discloses "detecting an object in a video" and "detecting a plurality of attributes of the object." Section II.B.2; Pet. 27–28 (citing Ex. 1003, 24–25, 50, 68–69, 71, 77–80, Figs. 4–8). Kellogg specifically discloses detecting objects, assigning an identifier to the object, tracking the detected objects, and storing information on the objects in visual memory. *See* Ex. 1003, 50, 68–69. Kellogg further discloses the detection of spatial and temporal attributes of an object, including type, location, duration, and times. *Id.* at 24–25, 71, 77–80, Figs. 4–8.

IPR2018-00138-25, p. 18, Final Written Decision

The PTAB has already decided that *Kellogg* discloses object and attribute detection that is independent from event determination

We disagree with Patent Owner. Kellogg discloses storing basic attribute information for an object, such as spatial, temporal, or spatiotemporal information.

Ex. 1003, 53–63.

that are people. *Id.* at 79. The system stores attribute information of objects, regardless of the user rule or query specified. *Id.* This allows the user to specify the query, and the visual memory system is able to retrieve objects and attributes based on the specified event. *Id.* That is, for any given monitored scene, the system collects and stores attributes regardless of a rule or query created by a user monitoring the scene.

IPR2018-00138-25, p. 18-19, Final Written Decision

The PTAB has already decided that *Kellogg* discloses object and attribute detection that is independent from event determination

Furthermore, we agree with Petitioner's Reply argument that Kellogg's "Approach Event" discloses the "independence-based claim elements." Pet. Reply. 11–14. Kellogg collects basic centroid attributes of two objects, such as spatial and temporal information. Ex. 1003, 62–63; *see* Pet. 12. As Petitioner notes, a query by a user to determine the distance between the two objects uses the spatial attribute information to calculate the distance between the two objects. Pet. 12. That is, the spatial and temporal attributes are detected regardless of a later created query by a user, *i.e.*, the detection of attributes is unaffected by the user rule.

IPR2018-00138-25, p. 19, Final Written Decision

"Attributes that are detected are <u>independent</u> of which event is identified" - Independence Arg. 2

Petitioners' Construction/ Position	Patent Owner's Construction/Position
Event detection process does not alter the attribute detection process.	"independent" means "the attributes are detected without regard to or knowledge of events or identification of events."

Board's Institution Construction
At this stage of the proceeding, we need not construe this claim limitation as Petitioner provides evidence and argument that Kellogg discloses this limitation under either claim interpretation. See Pet. 40–42.

Board's '661 FWD Construction
On the full record before us, we agree with Petitioner that the "independence-based claim elements" should be construed to require that "the detection of attributes is independent from, i.e., not affected by, the user rule that tasks the system." IPR2018-00138, FWD at 8.

Indexing does not turn attributes into events

Patent Owner further argues that Kellogg discloses "event-indexing," which merely references an already determined/detected event. PO Resp. 36–41. We are not persuaded by Patent Owner. Kellogg discloses "an indexing mechanism to quickly identify objects meeting sets of constraints." Ex. 1003, 64. We, however, are unable to discern the relevance of Kellogg's disclosure of indexing with the "independence-based claim elements." *See* Tr. 43:12–44:12. Kellogg discloses that the indexing "can improve retrieval performance." Ex. 1003, 64. Although indexing allow for the fast retrieval of predetermined objects and attributes, such as temporal or spatial, the indexing is not relevant to Kellogg's disclosure of the detection of attributes.

IPR2018-00138-25, p. 19-20, Final Written Decision

NEW USER RULE

“New User Rule”

Petitioners’ Construction/ Position	Patent Owner’s Construction/Position
A specified combination of a set of attributes for identifying an event	Plain and ordinary meaning, which is a “new a set of conditions such that when a defined event is detected it may trigger a response,”

Board’s Institution Construction
No explicit construction of “new user rule” is necessary

Board’s ’661 FWD Construction
Although Patent Owner argues that a rule requires more than a query that returns whether an event has occurred (PO Resp. 31–32; Sur-Reply 9–10), we agree with Petitioner that a “response” is not required. IPR2018-00138, FWD at 13.

Kellogg discloses providing a response

The user can specify for each region what types of objects are important, as shown in the right half of Figure 4-8. For example, the query region might return all objects, an indoor alarm region only people, and an outdoor alarm region both people and vehicles. Alarm regions default to the same type specification as the query region.

Associated with each alarm region is a delay specification that indicates how long

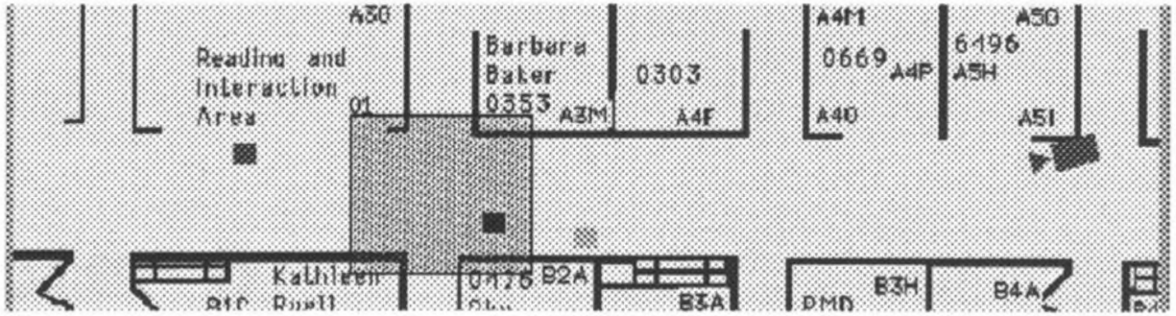


Figure 4-9: Graphical query results

an object must remain in that region before the system triggers an alarm. This lets

IPR2019-00311 Ex. 1003 (Kellogg) at 79-80

APPLYING THE NEW USER RULE TO ATTRIBUTES

The PTAB has held that Kellogg "Applies" user rules to attributes

’661 Patent Requires – "identifying an event of the object ***by applying*** the user rule to at least some of the plurality of attributes of the object"

The FWD of the ‘661 Patent (IPR2018-00138) found Kellogg meets this limitation. At. p. 16.

‘923 Patent Requires – "identifying an event of the object ... ***by applying*** the new user rule to the plurality of detected attributes,"

“Applying” – Independence Argument 1

Petitioners’ Construction/ Position	Patent Owner’s Construction/Position
“applying” would encompass any mechanism for analyzing the detected attributes to determine if they satisfy the user rule criteria, e.g., querying a database	Plain and ordinary meaning; when “applying the new user rule to the plurality of detected attributes” some level of analysis occurs that is greater than mere data retrieval

Board’s Institution Construction
No construction of the term is required

'923 Patent Disclosure

Although the video surveillance system of the invention draws on well-known computer vision techniques from the public domain, the inventive video surveillance system has several unique and novel features that are not currently available. For example, current video surveillance systems use large volumes of video imagery as the primary commodity of information interchange. The system of the invention uses video primitives as the primary commodity with representative video imagery being used as collateral evidence. The system of the invention can also be calibrated (manually, semi-automatically, or automatically) and thereafter automatically can infer video primitives from video imagery. The system can further analyze previously processed video without needing to reprocess completely the video. By analyzing previously processed video, the system can perform inference analysis based on previously recorded video primitives, which greatly improves the analysis speed of the computer system.

IPR2019-00311, Ex. 1001, 5:6-23

KELLOGG's Approach Event

```
Select p from Person
  during %times-1
  where p in
     (Select q from Person
       where q centroid within 3 of %spatiotemporal-spec
       during %times-2)
```

IPR2019-00311 Ex. 1003 (Kellogg) at 63

APPLYING THE NEW USER RULE TO ONLY THE ATTRIBUTES

“Applying the new user rule to only the plurality of detected attributes”

Petitioners’ Construction/ Position	Patent Owner’s Construction/Position
Ordinary Meaning, or Excludes coverage of systems that always reference an object hierarchy structure such as a tree structure that requires traversal of abstractions to apply the user rule	Any system that has the ability to search both attributes and abstractions, like in Kellogg—even if done at separate times—does not fall under the claim language of the ’923 patent.

Board’s Institution Construction
“Petitioner provides evidence and argument that the asserted prior art can ‘search only the attributes themselves and does not require traversing a tree structure of abstractions to search the detected attributes.”

‘923 Patent specification has no support for Avigilon’s “only” construction

In block **31**, one or more objects types of interests are identified in terms of video primitives or abstractions thereof. Examples of one or more objects include: an object; a person; a red object; two objects; two persons; and a vehicle.

IPR2019-00311, Ex. 1001, 8:16-19

In block **35**, one or more discriminators are identified by describing interactions between video primitives (or their abstractions), spatial areas of interest, and temporal attributes of interest. An interaction is determined for a combination of

IPR2019-00311, Ex. 1001, 8:50-54

‘923 Patent specification has no support for Avigilon’s “only” construction

Anytime after a video source has been processed according to the invention, video primitives for the source video are archived in block **43** of FIG. **4**. The video content can be reanalyzed with the additional embodiment in a relatively short time because only the video primitives are reviewed and because the video source is not reprocessed. This provides a great efficiency improvement over current state-of-the-art systems because processing video imagery data is extremely computationally expensive, whereas analyzing the small-sized video primitives abstracted from the video is extremely computationally cheap. As an example, the following event

IPR2019-00311, Ex. 1001, 14:61-15:4

Crystal Semiconductor Corp. v. TriTech Microelectronics Int'l, Inc., 246 F.3d 1336 (Fed. Cir. 2001)

- "'comprising' ***creates a presumption*** that the body of the claim is open…[and] that the claim does not exclude additional, unrecited elements.'" *Id.* at 1348.

- "Because claim 4 uses 'comprising,' it encompasses more than one clock ***unless the written description or the prosecution history clearly limits*** claim 4 to its recited elements." *Id.* at 1351.

Avigilon's Expert admits that adding abstractions does not avoid the claim

Q. Let's make it clear.

I build a system that does -- and I admit it does -- every single thing in the claims. It detects objects, it identifies attributes of objects, every limitation. It applies new user rules to attributes.

In addition to doing all of those things, it also applies new user rules to abstractions. Does that mean my product doesn't infringe?

A. I mean, what you're saying is I have a product that practices every element of the claim --

Q. Mm-hmm.

A. -- and then does something else for some unknown purpose?

Q. Correct.

A. Okay. I mean, it sounds like your product is at -- practicing every element of the claim.

THE COURT REPORTER: I'm sorry. Practice?

THE WITNESS: Practicing every element of the claim and, you know, the rest of that sounds irrelevant.

Bovik Deposition
IPR2019-00311, Ex. 1056, 207:15-208:15

Kellogg Searches Only The Attributes

In addition to the indices described above, each test also includes a "bucket" index. A bucket index simply maintains a list of all the objects stored in the visual memory. Since there is no overhead for the storage of complex index structure, a bucket index can achieve the highest update performance. A bucket index answers a query by retrieving all the objects in its list and checking them against the query specification.

As expected, the bucket indices achieve the highest performance for small numbers of objects. However, the temporal bucket cannot store much more than 100 updates,

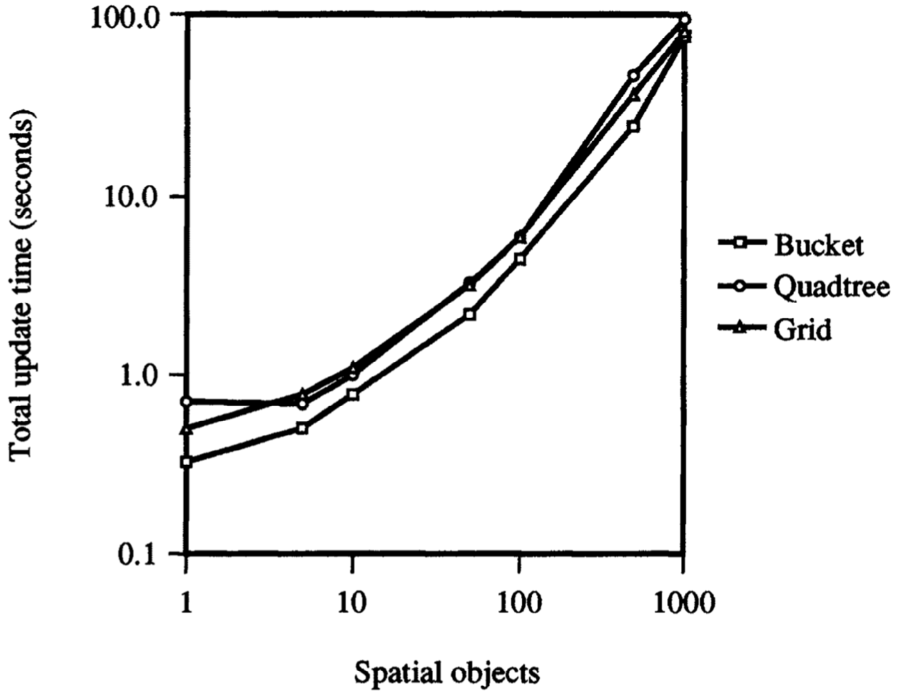


Figure 5-2: Spatial update performance

IPR2019-00311 Ex. 1003 (Kellogg) at 83-84

SINGLE CAMERA

Single Camera Limitations

[1.1] detecting an object in a video from a single camera;

[1.2] detecting a plurality of attributes of the object by analyzing the video from said single camera, the plurality of attributes including at least one of a physical attribute and a temporal attribute, each attribute representing a characteristic of the detected object;

'923 Patent discloses using multiple cameras, not single

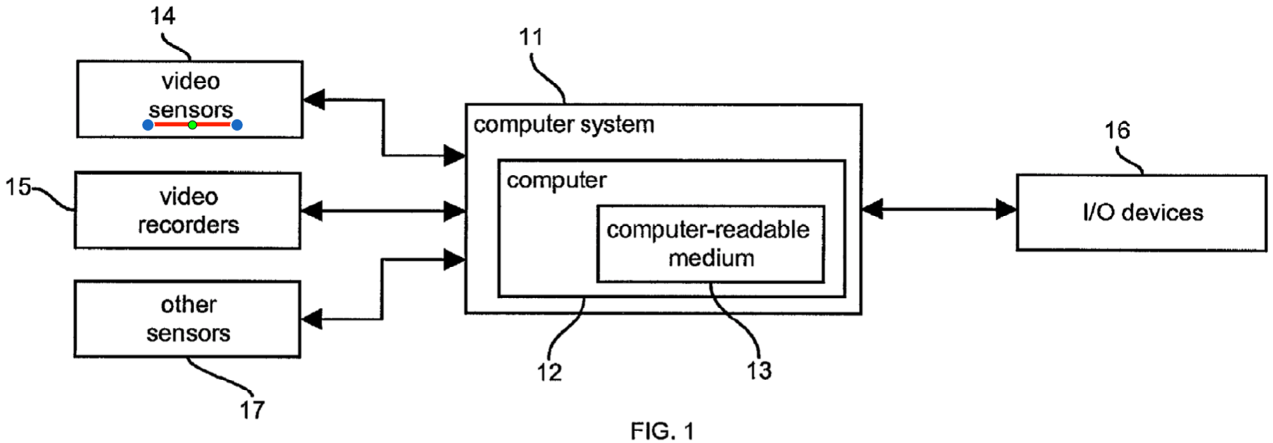


FIG. 1

The video sensors **14** provide source video to the computer system **11**. Each video sensor **14** can be coupled to the computer system **11** using, for example, a direct connection (e.g., a firewire digital camera interface) or a network. The video sensors **14** can exist prior to installation of the invention or can be installed as part of the invention. Examples of a video

IPR2019-00311, Ex. 1001, 6:3-8

In block **41**, the computer system **11** obtains source video from the video sensors **14** and/or the video recorders **15**.

IPR2019-00311, Ex. 1001, 9:23-24

For manual calibration, the operator provides to the computer system **11** the orientation and internal parameters for each of the video sensors **14** and the placement of each video sensor **14** with respect to the location. The computer system **11** can optionally maintain a map of the location, and the placement of the video sensors **14** can be indicated on the map. The map can be a two-dimensional or a three-dimen-

IPR2019-00311, Ex. 1001, 12:51-57

Avigilon's Expert's Testimony Re Single Camera

Q. At the time the 923 Patent was filed, was object detection from a single camera a unique new thing first described in this patent?

A. I mean, object detection had been around for a long time for one or more cameras. So I haven't stated that -- that, you know, the invention of Lipton is object detection from a single camera. It's a limitation of the claim is what it is.

Bovik Deposition
IPR2019-00311, Ex. 1056, 21:19 - 22:4

Avigilon's Expert's Testimony Re Single Camera

Q. So you wouldn't read this as a negative limitation to say, a system that does -- and we're just talking about this limitation -- object detection with video from a single camera is excluded from the claim because it has another camera doing something unrelated somewhere else; right?

A. Well, if there's another camera doing something unrelated somewhere else, it's completely unrelated to this claim as well.

Bovik Deposition
IPR2019-00311, Ex. 1056, 183:4-13

Avigilon’s Expert testified that Kellogg disclosed no problem with object detection

disagree. I find Dr. Grindon’s claims unsupported because there is no place in Kellogg that says that it is struggling with tracking objects or that objects were being occluded. Nor do I, or would any POSITA, view any aspect of Kellogg as indicating that it was a problem for the Kellogg reference. Thus, there would be no motivation to look to a reference like Brill.

Bovik Declaration
IPR2019-00311, Ex. 2019, 29-30

Kellogg teaches single camera object detection with more Detail than the ‘923 Patent

The input for the scene monitoring prototype comes from real-time processing of CCD camera images. This software, which tracks people walking in its field of view, was implemented by Tom Bannon and Tom O'Donnell in the Image Understanding Branch at the Texas Instruments Computer Science Laboratory. Using a calibrated internal model of its field of view, the software estimates the positions and heights of people and updates the visual memory a few times per second. This yields enough information to test the visual memory's performance and to provide interesting data for queries to retrieve.

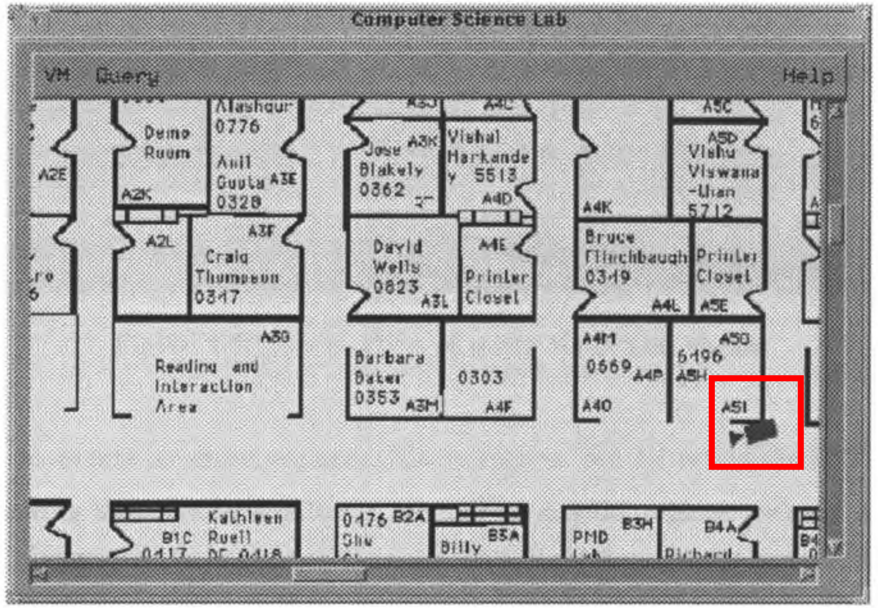


Figure 4-7: Graphical query interface viewing region

IPR2019-00311 Ex. 1003 (Kellogg) at 77-78

Kellogg teaches single camera object detection with more Detail than the ‘923 Patent

The user can specify for each region what types of objects are important, as shown in the right half of Figure 4-8. For example, the query region might return all objects, an indoor alarm region only people, and an outdoor alarm region both people and vehicles. Alarm regions default to the same type specification as the query region.

Associated with each alarm region is a delay specification that indicates how long

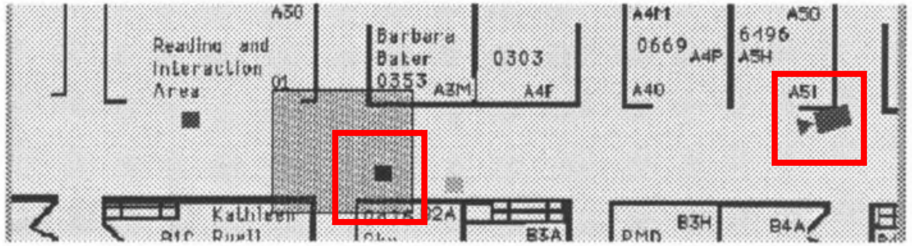


Figure 4-9: Graphical query results

an object must remain in that region before the system triggers an alarm. This lets

Once a query is fully specified, the results can be displayed in one of three ways: playback, event report, and trail trace. A playback steps through the retrieved information in temporal order, displaying moving blocks for moving objects and printing alarm information in another window. An event report textually describes alarms that were triggered. A trail trace displays blocks for all the retrieved information simultaneously and provides information about a certain object in response to a button press over its picture. Figure 4-9 shows part of a playback window, with one object inside an alarm region and two other objects also being monitored.

IPR2019-00311 Ex. 1003 (Kellogg)
at 79-80

Kellogg teaches single camera object detection with more Detail than the '923 Patent

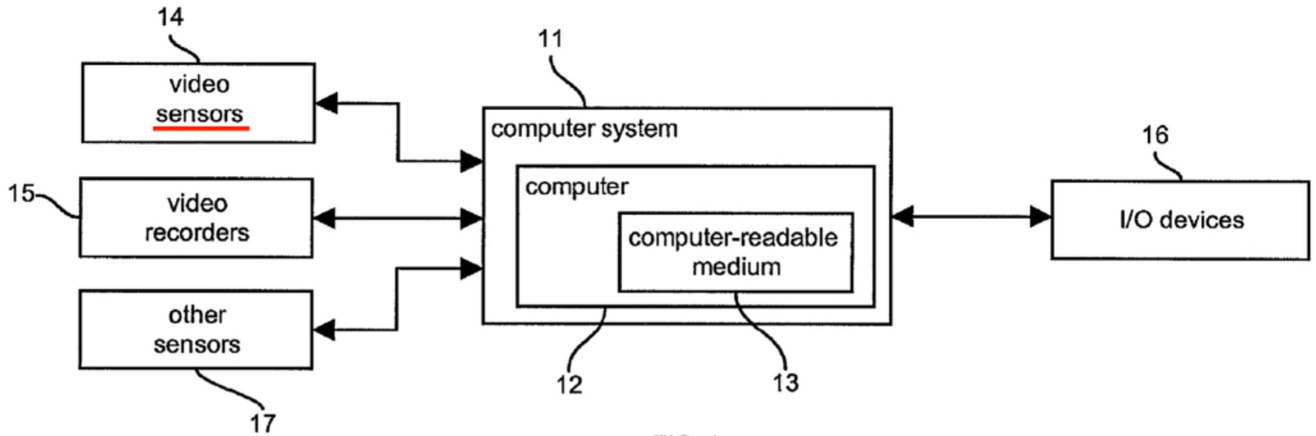


FIG. 1

'923 Patent

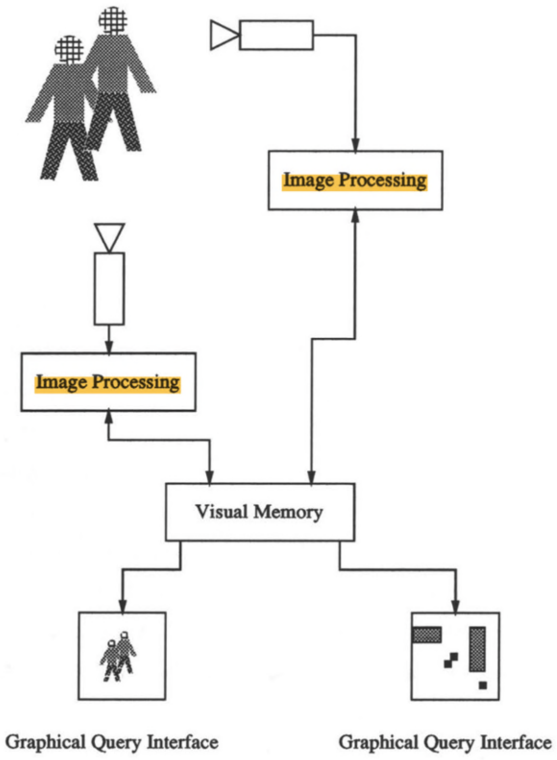


Figure 4-1: Scene monitoring prototype

KELLOGG

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Brill teaches the "Single Camera" limitations

Figure 1 shows a diagram of the AVS system. One or more "smart" cameras process the video stream to recognize events. The resulting event streams are sent to a Video Surveillance Shell (VSS), which integrates the information and displays it on a map. The VSS can also generate alarms based on the information in the event streams. In recent

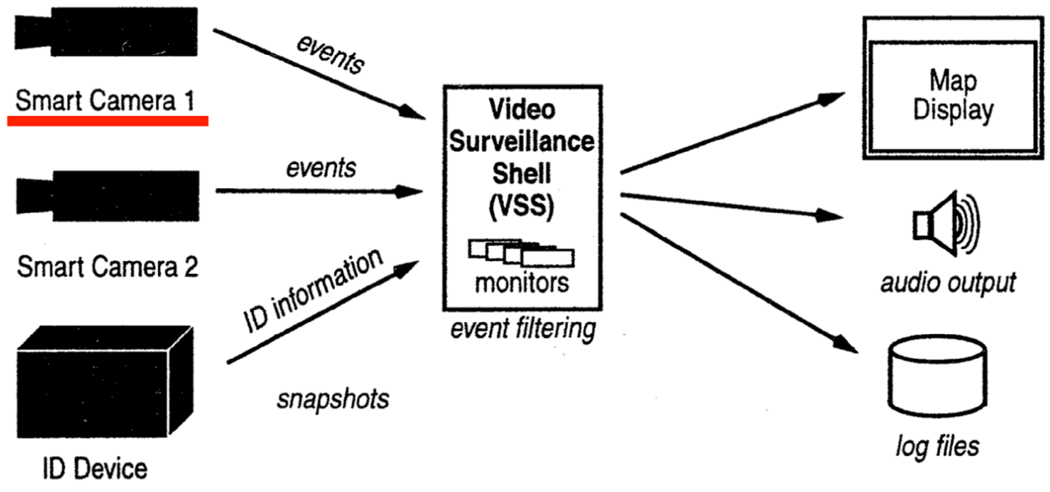


Figure 1: AVS system diagram

Brill Ex. 1004, 4-5

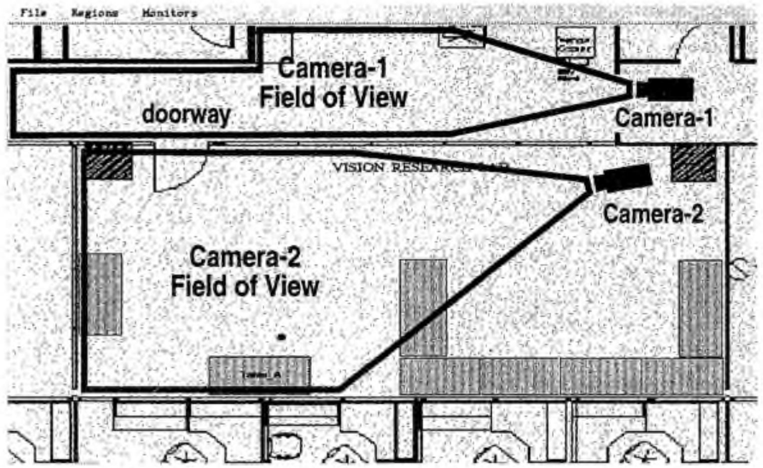


Figure 2: Multiple cameras with adjacent fields of view

2.3.6.2. Cloudy day. This is a 13 minute sequence in the same location as the previous sequence except it is a cloudy day. In this time span, 9 cars passed through the camera's field of view and all of them were detected by the system. There were a total of 2 people entering a car and 2 people exiting a car. The system successfully detected them all. Ad-

Brill Ex. 1004, 10

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

A POSA would combine Kellogg and Brill for several reasons

vehicle). *Brill* teaches that its "new approach involve[es] additional image differencing…[which] allows objects to be detected and tracked even when their images overlap the image of the car." Ex. 1004, 6. *Brill* specifically teaches a background-model based technique. *Id.*, 6-9. And a POSITA would have been highly motivated to combine the teachings of *Brill* with *Kellogg*'s monitoring system to solve the loss of tracking issue.

173. *Brill* teaches that even after an occlusion, the objects are reliably detected and tracked by relying on non-overlapping areas. *Id.*, 15. A POSITA would have been highly motivated to combine the teachings of *Brill* with *Kellogg*'s monitoring system to solve the occlusion issue and enhance detection and tracking of multiple moving objects.

174. The motivation to combine *Kellogg* and *Brill* is further evidenced by the fact that they were both related to the AVS systems developed by Texas Instruments. Indeed, at least one prior art publication confirms that the visual memory database of *Kellogg* was combined with the AVS system disclosed in *Brill*. Ex. 1007, ¶¶40-44.

IPR2019-00311, Ex. 1005 (Grindon Decl.), ¶¶ 170, 173, 174

The PTAB has already decided that a POSA would combine Kellogg and Brill

Kellogg's visual memory database." *Id.* (citing Ex. 2007 ¶ 105). We disagree with Patent Owner. Specifically, we find Petitioner is asserting Brill only for the functionality of "detecting an object in a video." *See* Pet. Reply 16; Pet. 58. As discussed above, Brill discloses a background-model-based technique for object detection. Ex. 1004, 6–9. Accordingly, we agree with Petitioner that a person with ordinary skill in the art would have been aware of video processing systems utilizing object detection, and a person with ordinary skill in the art would have combined Kellogg and Brill in order to "provide enhancements or achieve particular design objectives." Pet. 25.

IPR2018-00138-25, p. 27, Final Written Decision

There are no relevant "secondary considerations"

The prevailing systems at the time were focused on systems specific to a particular application that would detect known objects with known characteristics. *See id.* at 62. In fact, the prior art taught away from more generalized systems, which were complex and difficult to build. The '923 patent, however, provides a novel and inventive technique for event identification ***independent*** from the attribute and object detection. As a result, the '923 patent system allows "a developer [to] build features to apply to a variety of different contexts depending on an end-user's needs" and would not be limited to developing single-application systems. *Id.* Thus, the

Avigilon's Argument (IPR2019-00311, 47)

Q. So I would have to, as a designer, have knowledge of the implementation I want to make and the scenarios I want to cover to identify the attributes that need to be collected; am I correct?

A. I think, you know, a person of ordinary skill in the art would need some information about what they're applying this to. And I think this is -- this is something I would say just -- just about any patent involving image analysis methods.

Q. Right. So if I was going to develop a system for a grocery store, I might collect different attributes than if I was going to develop a system for -- a traffic analysis system?

A. Might be.

Q. Does the claim require that a system covered by the claim has to be able to operate in both the grocery store environment and in the traffic analysis environment?

A. I haven't read anything like that.

IPR2019-00311, Ex. 1056 (Bovik Dep.) 122:4-123:1

There are no "secondary considerations"

You can build an embodiment consistent with the claims of the 923 Patent to take a video of a car passing through a scene and extract attributes and objects about that car passing through the scene. Is that correct or incorrect?

A. It could be correct. I mean, I don't know the end application here, why you're doing this or anything. You know, and all that information could color any answer that anybody could give.

And, you know, if you're a person of ordinary skill in the art, how you would address that question that you asked me as a person of ordinary skill in the art trying to practice a patent, they would have to have an actual application in mind and understand the dimensions and the parameters of that application, and their hopeful -- you know, hopeful outcome and other things. A bunch of other stuff, too.

Q. Okay. So before attempting to embody the claims of the 923 Patent, an engineer would need to know the domain that is intended to be analyzed by the system?

A. I think that's true of any patent involving image analysis. So, sure.

IPR2019-00311, Ex. 1056 (Bovik Dep.) 92:21-93:23

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

VIDEO DEVICE

Kellogg teaches “a Video Device,” like the ‘923 Patent

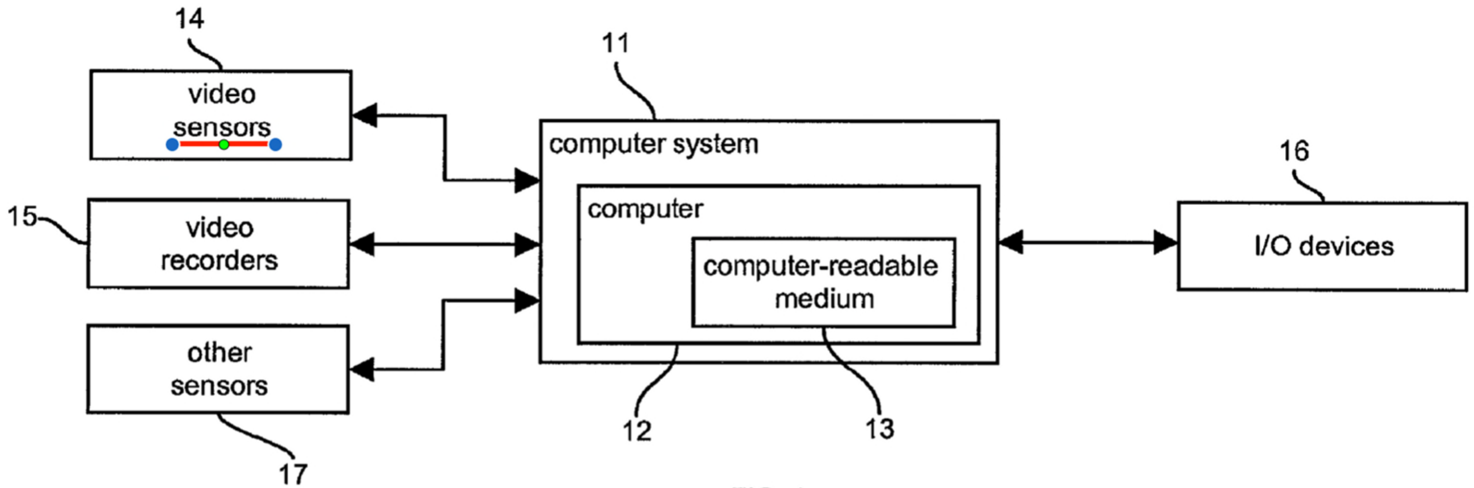


FIG. 1

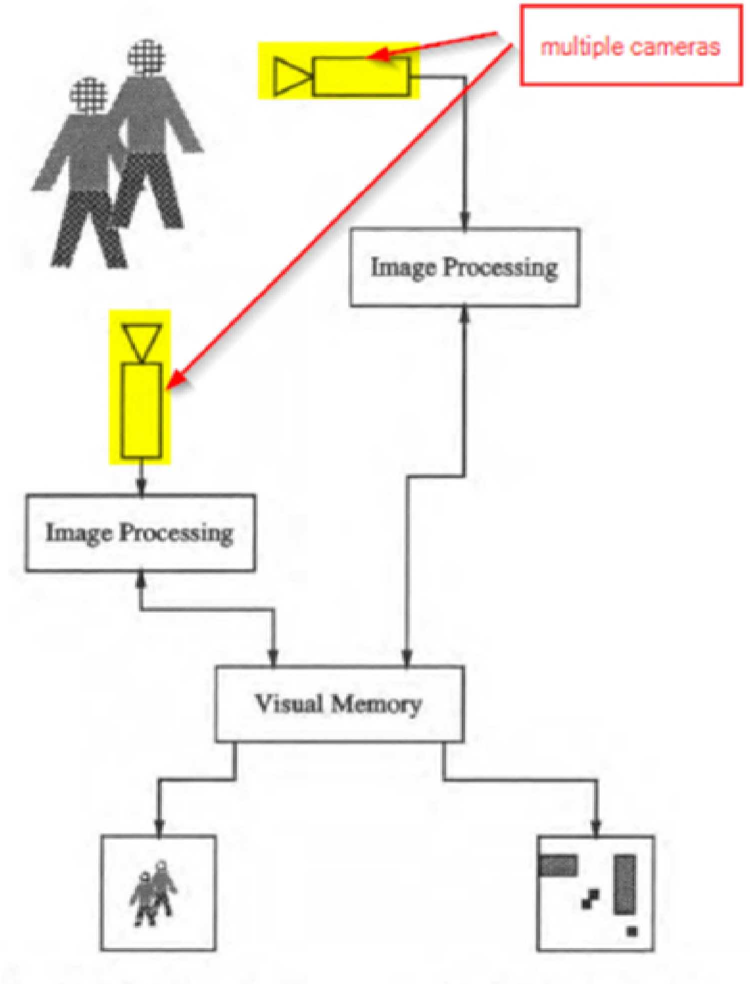


Figure 4-1: Scene monitoring prototype

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Kellogg teaches "a Video Device"

I mean, just to point out more of what Dr. Grindon said, I mean, he -- he points to figure 41, you know, and as I state, you know, figure 41 does show some video devices. I see video cameras here and a separate visual memory.

Bovik Deposition
IPR2019-00311, Ex. 1056, 148:4-9

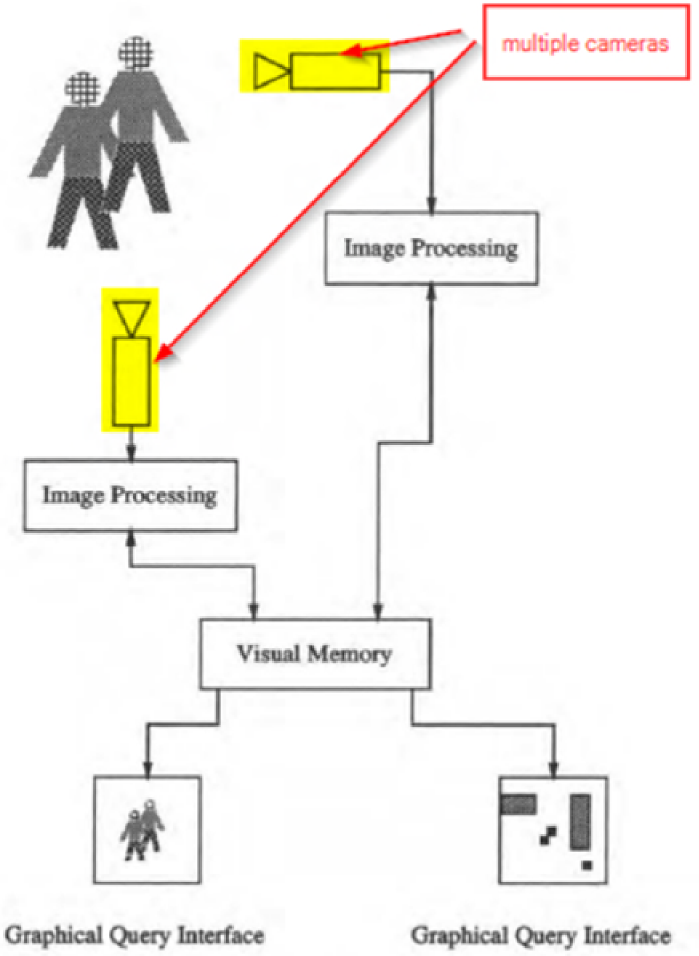


Figure 4-1: Scene monitoring prototype

DIMITROVA

Dimitrova teaches storing attributes and detecting events

Motion Recovery for Video Content Classification

NEVENKA DIMITROVA and FOROUZAN GOLSHANI
Arizona State University, Tempe

Like other types of digital information, video sequences must be classified based on the semantics of their contents. A more-precise and completer extraction of semantic information will result in a more-effective classification. The most-discernible difference between still images and moving pictures stems from movements and variations Thus, to go from the realm of still-image repositories to video databases, we must be able to deal with motion. Particularly, we need the ability to classify objects appearing in a video sequence based on their characteristics and features such as shape or color, as well as their movements By describing the movements that

IPR2019-00314 Ex. 1006
(Dimitrova) at 3

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Dimitrova

Motion is an essential feature of video sequences. By analyzing motion of objects we can extract information that is unique to the video sequences. In human and computer vision research there are theories about extracting motion information independently of recognizing objects. This gives us support for the idea of classifying sequences based on the motion information extracted from video sequences regardless of the level of recognition of the objects. For example, using the motion information we can not only submit queries like "retrieve all the video sequences in which there is a moving pedestrian and a car" but also queries that involve the exact position and trajectories of the car and the pedestrian.

IPR2019-00314 Ex. 1006
(Dimitrova) at 4

The PTAB has already analyzed Dimitrova

Dimitrova is directed to content retrieval based on video data for video on demand, automated surveillance systems, and other similar systems. Ex. 1003, 3–4. Dimitrova discloses that the distinction between still images and moving pictures are typically based in movements and variations, and, accordingly, the analysis of the motion of objects allows for the extraction of information that is unique to a video sequence. *Id.* at abstract, 4.

IPR2018-00140, p. 7, Final Written Decision

The PTAB has already analyzed Dimitrova

Dimitrova teaches that motion analysis of objects enables specific queries based on object data, such as object trajectories. *Id.* at 11. For example, Dimitrova states that a trajectory query to ""retrieve objects that have a motion trajectory whose point of origination is at the main gallery door and terminate at the Juan Miro's picture on the opposite wall"" provides information as to which object, such as a person, interacted or damaged the Juan Miro picture. *Id.*

IPR2018-00140, p. 7, Final Written Decision

ATTRIBUTES VS. EVENTS

ATTRIBUTES VS. EVENTS

- Avigilon makes several arguments all based on the faulty premise that the prior art detects events instead of attributes, see:

  - Dimitrova does not disclose “the plurality of attributes that are detected are independent of which event is identified” POR at 24.
  - Dimitrova does not disclose “identifying…applying the new user rule to only the plurality of detected attributes” POR at 16.

The PTAB has already decided that *Dimitrova* is not distinguished because it is like *Courtney.*

We disagree with Patent Owner. Specifically, we disagree with Patent Owner's arguments that an indexing system fails to disclose the "independence-based claim elements" for the reasons discussed above. Furthermore, we agree with Petitioner that Courtney "was distinguished because it detected a fixed set of 'events,' and users could only select among those 'events.'" Pet. Reply. 8–9. The

IPR2018-00140, p. 12, Final Written Decision

already selected as of interest. Ex. 2008, 4:62–67, 10:50–64. Accordingly, we find that Dimitrova is distinguished from Courtney because the Examiner found that Courtney discloses indexing events already selected as of interest, whereas Dimitrova discloses the detection or recognition of objects and activities based on attributes provided by a user.

IPR2018-00140, p. 12-13, Final Written Decision

OMV triplets record trajectory attributes just like '923 Patent

Example 4.1. A walking human may be represented as a moving object $a_1 = (o_1, m_1, v_1)$ where

$$o_1 = (category : human, convexHull : o_6 : skeleton : o_7, parts : \{head : o_2, torso : o_3\}),$$

$$m_1 = (trajectory : 2467332, activity : walking), \text{ and}$$

$$v_1 = (v\# : 234, firstFrame : 45, lastFrame : 485).$$

Ex. 1006 at 19, Example 4.1; Ex. 1038 ¶ 96.

A motion refers to any motion that can be automatically detected. Examples of a motion include: appearance of an object; disappearance of an object; a vertical movement of an object; a horizontal movement of an object; and a periodic movement of an object.

A salient motion refers to any motion that can be automatically detected and can be tracked for some period of time. Such a moving object exhibits apparently purposeful motion. Examples of a salient motion include: moving from one place to another; and moving to interact with another object.

A feature of a salient motion refers to a property of a salient motion. Examples of a feature of a salient motion include: a trajectory; a length of a trajectory in image space; an approxi-

IPR2019-00311, Ex. 1001, 7:37-49

Avigilon's Expert agrees the patent does not limit how attributes are stored

Q. Where does the patent say I can't use a data structure that stores an object, its classification and its motion?

A. The patent doesn't say that.

Bovik Deposition
IPR2019-00311, Ex. 1056, 114:15-18

Avigilon's Expert agrees that the Patent does not limit how attributes are stored

Q. Can the attributes be stored in an object-oriented database prior to being identified as an event?

A. I don't have an opinion one way or another. I mean, there could be an object-oriented database implicated in some way with an embodiment of this, but I haven't thought that. I certainly am not excluding it.

Bovik Deposition
IPR2019-00311, Ex. 1056, 85:15-23

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Dr. Grindon did not admit that Dimitrova merely detects events

Q And data is stored in Dimitrova in what is called an OMV triplet; right?

A Correct.

Grindon's Deposition
IPR2019-00311, Ex. 2018, 133:1-3

APPLYING THE NEW USER RULE TO ONLY THE ATTRIBUTES

Dimitrova permits searching only the attributes

The above functions allow for inexactness, and by default, they return results that are approximately similar to the precise answer. To make the operator exact, we use a higher-order operator that converts the query function in the desired manner, in this case, makes it exact. There are several types of these operators, e.g., exact, partial, and similar.

For making the retrieval exact, the symbol ! is placed in front of the query function. For example, "!*Agents*" returns only objects that have exactly the same motion description as the one given in M. Similarly, "!*Object_motion*" returns only motion descriptions of objects whose spatial characteristics (for example, exact bounding polygon, texture) match the spatial characteristics of a given object.

The # symbol placed in front of the query function is used for partial retrieval. Partial retrieval means that any of the motion or temporal characteristics of the given object should match. For example, "#*Agents*" will return all objects that match at least one of the motion descriptors.

IPR2019-00140, Ex. 1006, at 21

NEW USER RULE

The PTAB has already decided Dimitrova meets the user rule limitation

We disagree with Patent Owner. Patent Owner's argument is premised on a claim construction that we do not agree with and do not adopt. 138 Final Decision, 11–13. Rather, we construe the limitation "user rules that define an event" to mean "a set of attributes and/or other parameters for identifying an event." *Id.* Dimitrova discloses that a user provides a set of attributes and the system returns objects and activities based accordingly. Ex. 1003, 19. As such, we agree with Petitioner that Dimitrova discloses "user rules that define events." Pet. 29 (citing Ex. 1003, 20, 25); Ex. 1003, 19.

IPR2018-00140, p. 13, Final Written Decision

A POSA would combine Dimitrova and Brill for several reasons

1004, 6-9. *Brill* teaches improvements to object tracking so that tracking is not lost when a person's image overlaps that of another object, such as a car. In some systems this situation would cause the object to appear to merge resulting in the loss of tracking until the person walks away from the other object. *Id.* This problem would have been within the knowledge of a POSITA who employed *Dimitrova*'s system to monitor multiple objects. Take the scene including a moving toy and two cups, for example, in *Dimitrova*. Ex. 1006, Fig. 5.

182. *Brill* teaches that even after an occlusion, the objects are reliably detected and tracked by relying on non-overlapping areas. *Id.*, 15. A POSITA would have been highly motivated to combine the teachings of *Brill* with *Dimitrova*'s monitoring system to solve the occlusion issue and enhance detection and tracking of multiple moving objects.

183. Indeed, *Dimitrova* also signals to a POSITA that "[m]ore-sophisticated" detection algorithms would be beneficial. Ex. 1006, 32. Moreover, as both references teach a computer-based system employing video cameras, a POSITA would be able to predictably combine the teachings of the two references, without requiring extensive modification to the overall system. *Brill* would be an obvious source because, like *Dimitrova*, *Brill* is directed to video surveillance systems using object detection methods, attribute detection methods, and querying mechanisms. *See, e.g.*, Ex. 1006, 8-20, 25; Ex. 1004, 6-9, 12-14. Similar to *Dimitrova*'s system, *Brill's* system can detect multiple objects and their interactions in a single video scene. *See, e.g.*, Ex. 1006, 25; Ex. 1004, 6-9, 12-14.

IPR2019-00311, Ex. 1005 (Grindon Decl.), ¶¶ 178, 182, 183

The PTAB has already decided that a POSA would combine Dimitrova and Brill

We agree with Petitioner's arguments. Notwithstanding Patent Owner's arguments, which we address below, we determine Petitioner has demonstrated by a preponderance of the evidence that independent claim 1 of the '661 patent would have been obvious over Dimitrova and Brill. Petitioner provides a similar analysis for claims 2–32. *Id.* at 58–71. We agree with Petitioner's arguments for claims 2–32, and we similarly determine Petitioner has demonstrated by a preponderance of the evidence that claims 2–32 of the '661 patent would have been obvious over Dimitrova and Brill. *Id.*

IPR2018-00140, p. 17, Final Written Decision

Brill's "Actions" Show A Response

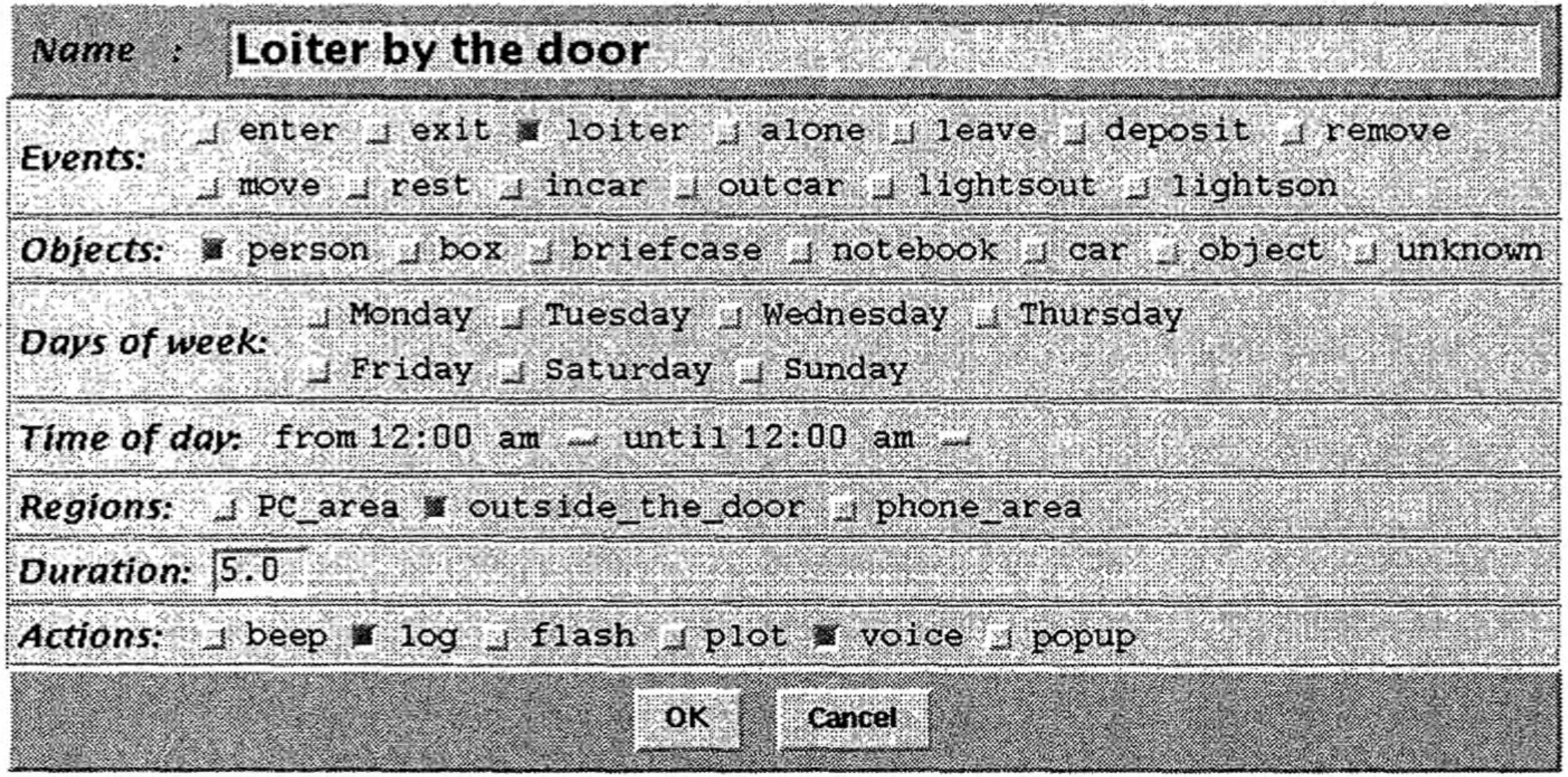


Figure 11: Selecting a type of simple event

PUBLICATION

Kellogg is a printed publication

Visual Memory

by

Christopher James Kellogg

IPR2019-00311, Ex. 1003 (Kellogg) p. 1

Barton MIT Libraries' Catalog

008	930928s1993 xx a b 000 0 eng d
035	|a MITb10656176
035	|a (OCoLC)28904769
035	|a GLIS00656176
040	|a MYG |c MYG
099	|a Thesis E.E. 1993 M.S. |a Thesis E.E. 1993 M.S. Mfch
099	|a Thesis |a E.E. |a 1993 |a M.S.
1001	|a Kollogg, Christopher James.
24510	|a Visual memory / |c by Christopher James Kellogg.
260	|c c1993.

IPR2019-00311, Ex. 1054 (Zimmerman Decl.) p. 9, Ex. B (Kellogg MIT MARC record)

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

MIT provided personal knowledge Kellogg was published

I, Katherine Zimmerman state and declare as follows:

1. I am a Scholarly Communications and Licensing Librarian at the Massachusetts Institute of Technology ("MIT") Libraries, 105 Broadway, Building NE36, Suite 6101, Cambridge, Massachusetts 02142.

8. The cover page of the original printed copy of the Kellogg Thesis has an MIT Libraries date stamp of "JUL 09 1993" indicating that the MIT Libraries received this thesis on July 9, 1993.

10. The Kellogg Thesis has a cataloging date of September 28, 1993 (shown as 930928 in line number 008 of Exhibit B), indicating it was entered into OCLC, the national bibliographic utility in which MIT does its cataloging, on September 28, 1993.

16. After being cataloged in OCLC, a thesis undergoes a process of being labeled and moved to a shelf of the MIT Libraries. This process typically takes two to four weeks. According to MIT's normal business practice at the time, the Kellogg Thesis would have been displayed on a shelf of the MIT Libraries no later than December 28, 1993.

IPR2019-00311, Ex. 1054
(Zimmerman Decl.) pp. 1-4

Brill is a printed publication

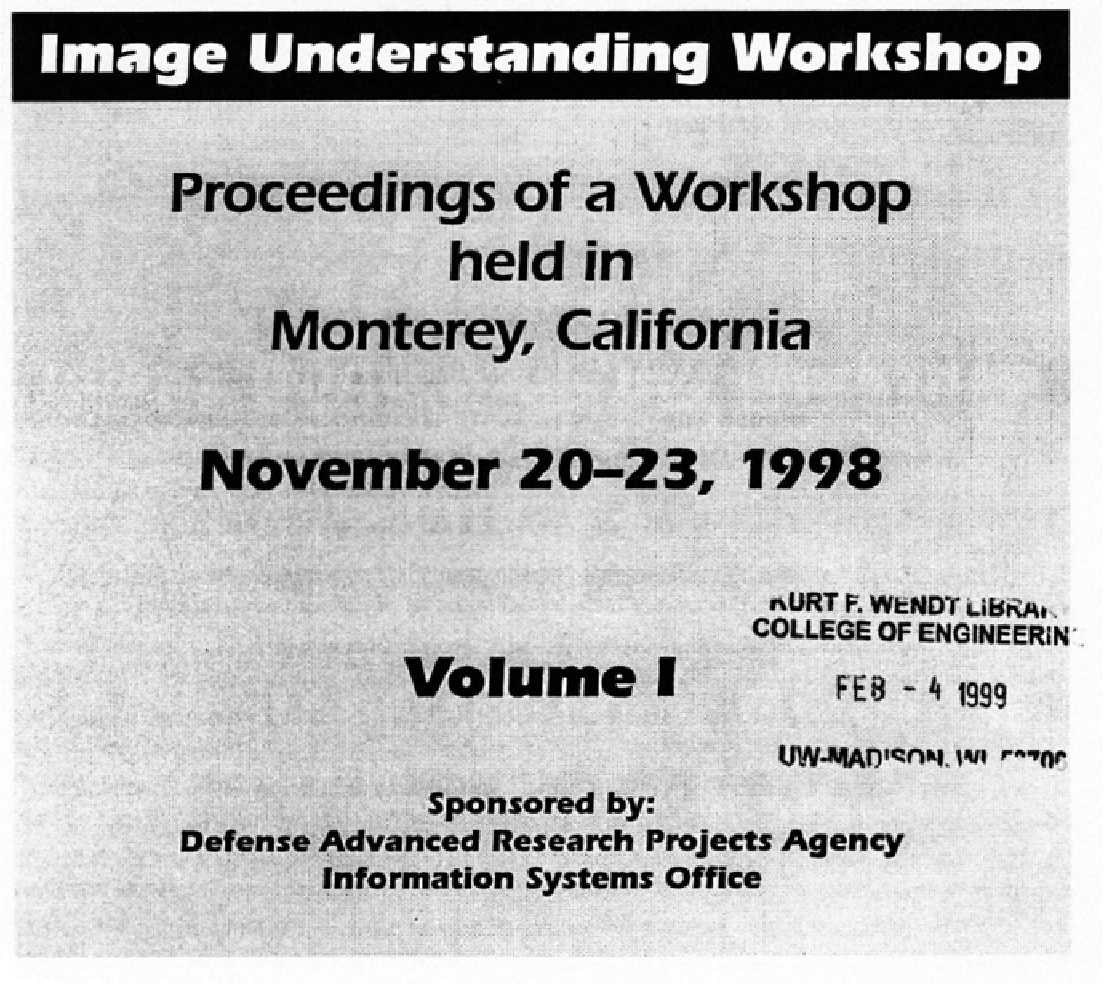


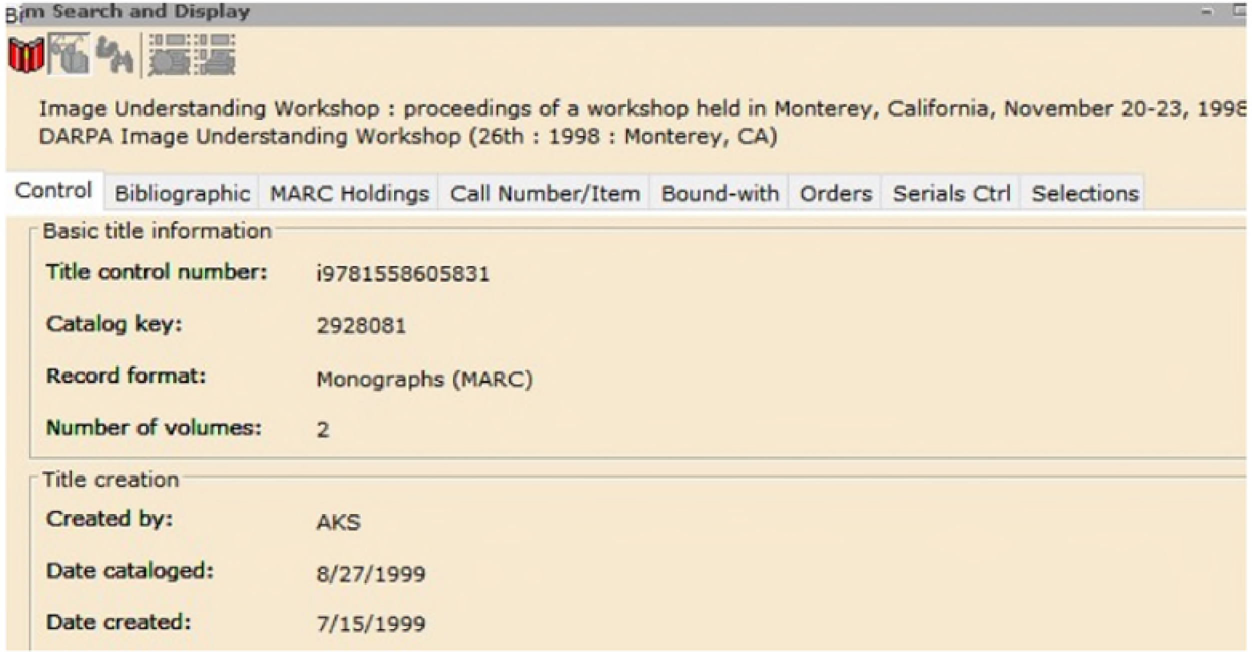
Image Understanding Workshop

Proceedings of a Workshop
held in
Monterey, California

November 20–23, 1998

Volume I

Sponsored by:
Defense Advanced Research Projects Agency
Information Systems Office



IPR2019-00311, Ex. 1049 (Kasik Decl.) p. 39 (Virginia MARC record)

IPR2019-00311, Ex. 1053 (Watters Decl.) p. 3 (Wisconsin stamped copy)

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Wisconsin and Virginia provided personal knowledge Brill was published

I, Rachel J. Watters, am a librarian, and the Director of Wisconsin TechSearch ("WTS"), located at 728 State Street, Madison, Wisconsin, 53706. WTS is an interlibrary loan department at the University of Wisconsin-Madison. I have worked as

Based on the information in Exhibit A, it is clear that the volume was received by the library on or before February 4, 1999, catalogued and available to library patrons within a few days or at most 2 to 3 weeks after February 4, 1999.

IPR2019-00311, Ex. 1053 (Watters Decl.) pp. 1-2

I, Bryan Patrick Kasik, state and declare as follows:

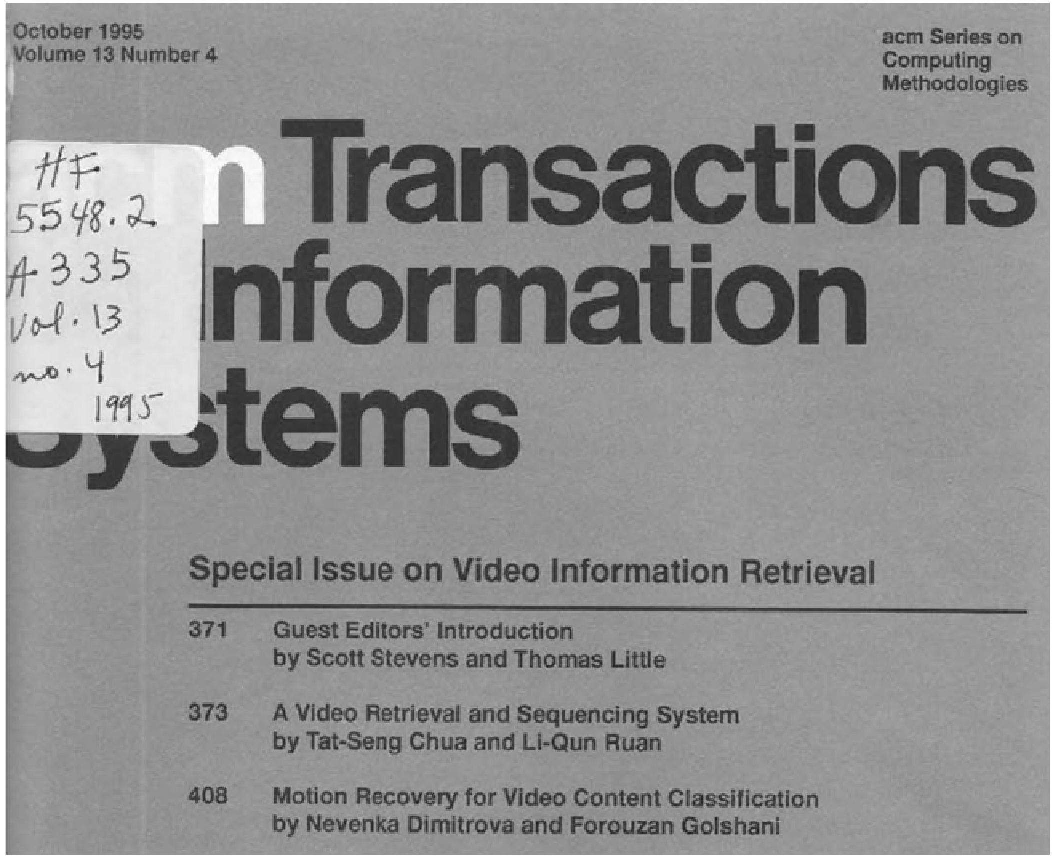
1. I am a Reference Librarian with the University of Virginia ("UVA") Library, 160 McCormick Road, Charlottesville, Virginia 22904.

7. According to the record in Exhibit B, Brill was ordered and entered into the UVA Library system on July 15, 1999, then assigned a call number on August 27, 1999.

8. After receiving a call number at the library, Brill was labeled and moved to the stacks. This process at the UVA Library typically takes a few days depending on the volume of books being processed. According to UVA Library's normal business practice, Brill would have been moved to the stacks and available to the public within a few days of August 27, 1999.

IPR2019-00311, Ex. 1049 (Kasik Decl.) pp. 2-3

Dimitrova is a printed publication

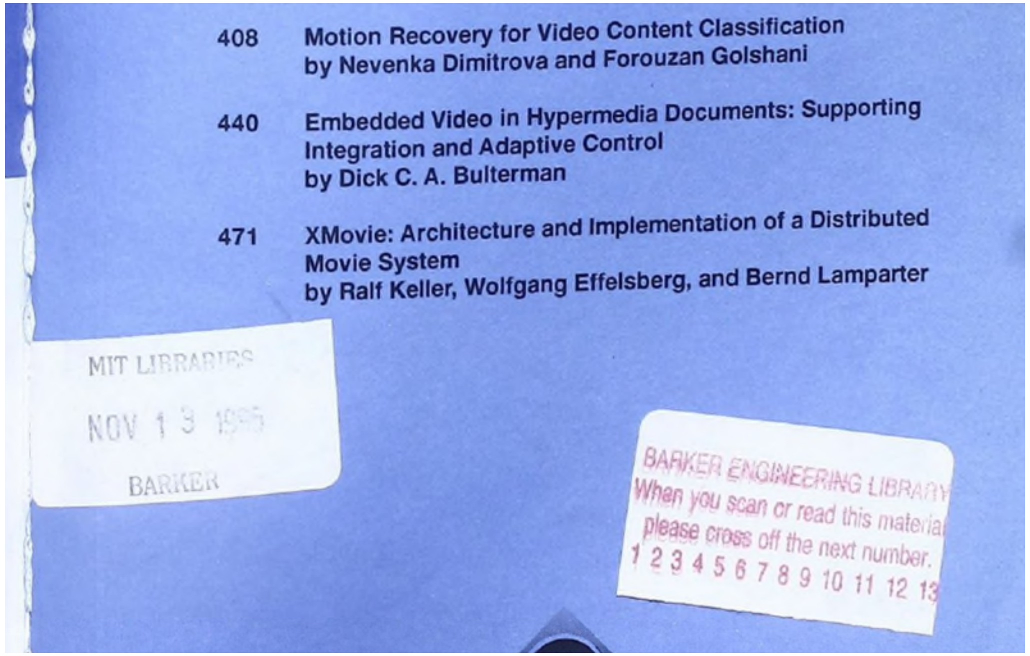

October 1995
Volume 13 Number 4

acm Series on Computing Methodologies

HF 5548.2 A335 Vol. 13 no. 4 1995

...n Transactions ...nformation ...ystems

Special Issue on Video Information Retrieval



IPR2019-00314, Ex. 1055 (Zimmerman Decl.) p. 10 (MIT stamped copy)

ACM Transactions on Information Systems (ISSN 1046-8188) is published 4 times a year in January, April, July, and October by the Association for Computing Machinery, Inc., 1515 Broadway, New York, NY 10036. Second-class postage paid at New York, NY 10001, and at additional mailing offices. Postmaster: Send address changes to Transactions on Information Systems, ACM, 1515 Broadway, New York, NY 10036.

IPR2019-00314, Ex. 1006 (Dimitrova) pp. 1-2

MIT provided personal knowledge Dimitrova was published

I, Katherine Zimmerman, state and declare as follows:

1. I am a Scholarly Communications and Licensing Librarian at the Massachusetts Institute of Technology ("MIT") Libraries, 105 Broadway, Building NE36, Suite 6101, Cambridge, Massachusetts 02142.

7. The October 1995 Issue has an MIT Libraries date stamp of "NOV 13 1995," indicating that the MIT Libraries received the issue on November 13, 1995.

8. After a serials issue receives a date stamp, it undergoes a process of being labeled and moved to a shelf of the MIT Libraries. Based on current MIT Libraries practice, this process typically takes one to two weeks. According to the MIT Libraries' current normal business practice, the October 1995 Issue would have been displayed on a shelf of the MIT Libraries no later than November 27, 1995.

IPR2019-00314, Ex. 1055 (Zimmerman Decl.) pp. 1-2

DEMONSTRATIVE EXHIBIT – NOT EVIDENCE

Attorney argument ***cannot*** defeat evidence of publication

- "It is well established that such bare attorney arguments cannot take the place of objective evidence and, thus, we accord them little evidentiary weight. *In re Payne*, 606 F.2d 303, 315 (CCPA 1979); *In re Pearson*, 494 F.2d 1399, 1405(CCPA 1974)('Attorney's argument in a brief cannot take the place of evidence')."
- "Even if we set aside Dr. Robinson's Declaration, however, we are persuaded that the indicia of publication on the face of Varenna 2012 are sufficient to establish that Varenna 2012 was 'sufficiently accessible to the public interested in the art' and 'disseminated or otherwise made available" to the interested public before the critical date, and consequently, a printed publication.' *Blue Calypso, LLC v. Groupon, Inc.*, 815 F.3d 1331, 1348 (Fed. Cir. 2016)."

END