UNITED STATES PATENT AND TRADEMARK OFFICE

---

BEFORE THE PATENT TRIAL AND APPEAL BOARD

---

Canon Inc., Canon U.S.A., Inc., and Axis Communications AB,

Petitioner

v.

Avigilon Fortress Corporation,

Patent Owner

---

Case: Unassigned

U.S. Patent No. 7,932,923
Issue Date: April 26, 2011

Title: Video Surveillance System Employing Video Primitives

---

# PETITION FOR *INTER PARTES* REVIEW

## TABLE OF CONTENTS

# LIST OF EXHIBITS

| | |
|---|---|
| Ex. 1001 | U.S. Patent No. 7,932,923 ("the '923 Patent") |
| Ex. 1002 | Prosecution History of the '923 Patent |
| Ex. 1003 | "Visual Memory" by Christopher James Kellogg ("*Kellogg*") |
| Ex. 1004 | "Event Recognition and Reliability Improvements for the Autonomous Video Surveillance System" by Frank Brill et al. ("*Brill*") |
| Ex. 1005 | Declaration of John R. Grindon, D.Sc. |
| Ex. 1006 | "Motion Recovery for Video Content Classification" by N. Dimitrova et al. ("*Dimitrova*") |
| Ex. 1007 | Declaration of Emily R. Florio |
| Ex. 1008 | February 29, 2012 Request for *inter partes* Reexamination of the '923 Patent |
| Ex. 1009 | May 23, 2012 Order Granting/Denying Request for *inter partes* Reexamination of the '923 Patent |
| Ex. 1010 | May 23, 2012 Office Action in *inter partes* Reexamination of the '923 Patent |
| Ex. 1011 | August 27, 2012 Amendment and Reply in *inter partes* Reexamination of the '923 Patent |
| Ex. 1012 | February 13, 2013 Decision Granting Petition to Terminate *inter partes* Reexamination Proceeding of the '923 Patent |
| Ex. 1013 | May 23, 2013 Attachment to Request for *ex parte* Reexamination of the '923 Patent |
| Ex. 1014 | June 17, 2013 Order Granting/Denying Request for *ex parte* Reexamination of the '923 Patent |
| Ex. 1015 | August 30, 2013 Office Action in *ex parte* Reexamination of the '923 Patent |

| | |
|---|---|
| Ex. 1016 | October 30, 2013 Amendment and Reply in *ex parte* Reexamination of the ’923 Patent |
| Ex. 1017 | April 4, 2014 Final Office Action in *ex parte* Reexamination of the ’923 Patent |
| Ex. 1018 | April 16, 2014 Amendment and Reply in *ex parte* Reexamination of the ’923 Patent |
| Ex. 1019 | April 30, 2014 Notice of Intent to Issue *ex parte* Reexamination Certificate of the ’923 Patent |
| Ex. 1020 | May 21, 2014 *ex parte* Reexamination Certificate of the ’923 Patent |
| Ex. 1021 | U.S. Patent No. 5,969,755 to Courtney (*“Courtney”*) |
| Ex. 1022 | “Object-Oriented Conceptual Modeling of Video Data” by Young Francis Day et al., (“*Day-I*”) |
| Ex. 1023 | Declaration of Christopher James Bailey-Kellogg in IPR2018-00138 and IPR2018-00140 |
| Ex. 1024 | Applicant Response of June 11, 2012 in *inter partes* Reexamination of U.S. Patent No. 7,868,912 (“the ’912 Patent”) |
| Ex. 1025 | Decision Granting Petition to Terminate *inter partes* Reexamination of the ’912 Patent |
| Ex. 1026 | May 24, 2013 Attachment to Request for *ex parte* Reexamination of the ’912 Patent |
| Ex. 1027 | June 20, 2013 Order Granting/Denying Request for *ex parte* Reexamination of the ’912 Patent |
| Ex. 1028 | August 30, 2013 Office Action in *ex parte* Reexamination of the ’912 Patent |
| Ex. 1029 | October 30, 2013 Amendment and Reply in *ex parte* Reexamination of the ’912 Patent |
| Ex. 1030 | March 27, 2014 Final Office Action in *ex parte* Reexamination of the ’912 Patent |

| | |
|---|---|
| Ex. 1031 | April 16, 2014 Amend and Reply in *ex parte* Reexamination of the ’912 Patent |
| Ex. 1032 | Notice of Intent to Issue *ex parte* Reexamination Certificate of the ’912 Patent |
| Ex. 1033 | Declaration of Kenneth A. Zeger for the *ex parte* Reexamination of the ’923 Patent |
| Ex. 1034 | U.S. Patent No. 7,868,912 |
| Ex. 1035 | U.S. Patent No. 8,564,661 |
| Ex. 1036 | Curriculum Vitae of John R. Grindon, D.Sc. |
| Ex. 1037 | Merriam-Webster’s Collegiate Dictionary |
| Ex. 1038 | Declaration of Dr. Alan Bovik in IPR2018-00138 and IPR2018-00140 |
| Ex. 1039 | The New York Times October 2, 1999, Pro Basketball; McHale and Thompson Enter Hall of Fame with 3 Others, by Mike Wise |
| Ex. 1040 | Specification of U.S. Serial No. 09/987,707 |

## I. INTRODUCTION

Canon Inc., Canon U.S.A., Inc. ("Canon") and Axis Communications AB ("Axis") (collectively "Petitioner"), request *inter partes* review ("IPR") of claims 1-41 of U.S. Patent No. 7,932,923 B2 and 7,932,923 C1 ("the '923 patent," Ex. 1001), owned by Avigilon Fortress Corporation ("Avigilon").

The goal of the '923 patent system is to reduce the processing required to identify and retrieve desired portions of video data. The challenge is that video data is voluminous, thus time consuming to process. The '923 patent's solution is to process the video once to collect information identifying objects and activities present in the video. Later, a user may define scenarios of interest, which are used to search the collected data and identify relevant video.

The '923 patent describes its solution in generally functional terms. For example, the patent assumes that a person of ordinary skill in the art ("POSITA") knows how to detect objects, identify object attributes, and search those attributes to identify the desired portions of the video, *i.e.*, events. Accordingly, these features are not disclosed in any technical detail. Rather than relying on technical distinctions, Avigilon distinguished its invention over the prior art by arguing that it processes video to detect objects and activities (attribute detection) "independently" from defining and searching for scenarios of interest (event detection).

The prior art relied on in this Petition discloses the so-called "independence" that Avigilon touts as the hallmark of its invention. Indeed, the Board has already affirmed this in the IPR Institution Decision of a related patent. IPR2018-00138, Paper No. 8, 15-16 (June 1, 2018).

As described in detail below, Petitioner has demonstrated by a preponderance of the evidence that each of the challenged claims is unpatentable. Therefore, the Board should also institute trial here.

## II. GROUNDS FOR STANDING

Petitioner certifies the '923 patent is available for IPR and Petitioner is not barred or estopped from requesting IPR of the '923 patent and challenging claims 1-41 on the grounds identified in this Petition.

## III. IDENTIFICATION OF CHALLENGES

### A. Claims for Which Review is Requested

Petitioner requests IPR under 35 U.S.C. § 311 of claims 1-41 of the '923 patent.

### B. The Prior Art and Specific Grounds on Which the Challenge to the Claims is Based

The specific statutory grounds, claims challenged and the prior art relied upon for each ground are:

| **Grounds for Challenged Claims 1-41** | |
|---|---|
| Ground 1 | Claims 1-41 are anticipated under pre-AIA 35 U.S.C. § 102(b) by "Visual Memory" by Christopher James Kellogg ("*Kellogg*," Ex. 1003) |
| Ground 2 | Claims 1-41 are obvious under pre-AIA 35 U.S.C. § 103(a) over *Kellogg* in view of "Event Recognition and Reliability Improvements for the Autonomous Video Surveillance System" by Frank Brill et al. ("*Brill*," Ex. 1004) |

The '923 patent issued April 26, 2011, from U.S. Application No. 12/569,116, filed September 29, 2009. Ex. 1001, 1. The '923 patent claims priority through a series of continuation-in-part applications to U.S. Application No. 09/694,712, filed October 24, 2000. *Id*. Therefore, the earliest possible priority date for the patent is October 24, 2000.[1]

*Kellogg* was publicly accessible at the Massachusetts Institute of Technology library in September 1993. Ex. 1003, 1; Ex. 1007, ¶¶22-29; Ex. 1023. *Kellogg* is prior art under pre-AIA 35 U.S.C. § 102(b).

[1] Petitioner does not concede that October 24, 2000 is the effective filing date of the '923 patent and reserves the right to challenge the priority claim.

*Brill* was published in the Proceedings of Image Understanding Workshop, Vol. 1, published December 1998. Ex. 1004, 1; Ex. 1007, ¶¶30-33. *Brill* is prior art under pre-AIA 35 U.S.C. § 102(b).

### C. How the Challenged Claims are Construed

Section V provides the construction of the challenged claims 1-41.

### D. How the Challenged Claims are Unpatentable

Section VIII describes how the challenged claims 1-41 are unpatentable in view of the prior art.

### E. Supporting Evidence

Exhibits supporting this Petition are attached. Ex. 1005 is a Declaration of John R. Grindon, D.Sc. under 37 C.F.R. § 1.68. Section VIII describes the relevance of the evidence to the challenged claims, including an identification of the specific portions of the evidence supporting the challenges.

## IV. THE ’923 PATENT

### A. Overview

The ’923 patent relates to a video surveillance system that extracts video “primitives” or “attributes” and determines the occurrence of an “event” based on the attributes. Ex. 1001, Abstract. The ’923 patent states that its detection and storage of attributes is advantageous over prior systems that searched raw video because reduces the amount of data to be stored or processed. *Id.*, 2:29-33. Examples of attributes include 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 scene change, etc. *Id.*, 7:8-12.

Event discriminators can be defined using objects, spatial and temporal attributes. *Id.*, 4:64-5:1. For example, a "loitering event" is defined as: (1) a "person" object, (2) in the "ATM" video space, (3) for "longer than 15 minutes," and (4) "between 10:00 p.m. and 6:00 a.m." *Id.*, 5:1-5.

The '923 patent does not describe the technical aspects of its system. Ex. 1005, ¶¶50-51. While it states that objects are detected (*id.* 9:30-44) and then attributes of the objects are detected (*id.* 10:49-52), it simply describes using "any" motion detection algorithm to detect objects. *Id.*, 9:33-35, 9:39-41. And it assumes that such detecting techniques are known in the art. *Id.*, 10:11-22, 10:27-30, 10:39-41, 10:44-47.

Similarly, the patent merely lists examples of attributes without describing how to identify those attributes. *Id.* 7:8-8:15, 10:49-51. Because all attributes are exemplary, the patent does not specify any particular set of attributes that are necessary to practice the invention. Ex. 1005, ¶¶53-57. As a result, the patent does not teach how to select attributes to identify any arbitrary set of events. *Id.*

## B. Prosecution History

### 1. Prosecution

During prosecution, the claims were found obvious over US 7,653,635 and US 6,721,454. Ex. 1002, 160-163. In response, the patentee amended the claims to recite selecting of a new user rule after detecting the plurality of attributes (or storing the detected attributes). *Id.*, 118-129.

After a first examiner interview, the claims were further narrowed to recite that the plurality of attributes include at least one of a physical attribute and a temporal attribute. *Id.*, 98-110. After a second interview, the claims were further narrowed to require video "from a single camera" to distinguish U.S. 2003/0023612. *Id.*, 78-93. The amended claims were allowed and matured into claims 1-41. *Id.*, 65-71.

### 2. *Inter Partes* and *Ex Parte* Reexaminations

The '923 patent was challenged in an *inter partes* reexamination, Control No. 95/001,914, by Bosch Security Systems, Inc. Ex. 1008, 1. The Patent Office instituted the reexamination on six grounds. Ex. 1009, ¶5; Ex. 1010, ¶¶3-10. Before any action by the examiner, the patentee and Bosch settled their dispute, and the reexamination was terminated. Ex. 1012, 4-5.

Later, the '923 patent was challenged in an *ex parte* reexamination, Control No. 90/012,876. Ex. 1013, 1. The Patent Office instituted the reexamination on multiple grounds. Ex. 1015, ¶¶10-21. The patentee filed an amendment replacing claims 1-41 with new claims 42-81. Ex. 1016, 2-21.

Specifically, claims 55-58 included a new limitation requiring that the step of applying the "new user rule" comprised applying the rule to ***only*** the plurality of ***detected attributes***. *Id.*, 8-10. And the patentee distinguished claims 55-58 from the prior art reference *Day-I* (Ex. 1022), arguing that "the queries of *Day-I* are not applied to the attributes stored in the VSDG alone but are applied to object-oriented abstractions." Ex. 1016, 78-79. But the patentee's interpretation of the new limitation lacks support from its citation to the patent. Ex. 1033, ¶30 (citing '707 application (Ex. 1040), ¶148 ("[t]he 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.")(emphasis in original); *see also* Ex. 1018, 7. In that context, the word "only" is used for excluding reprocessing source video as opposed to excluding something other than video primitives, such as abstractions. Ex. 1005, ¶69.

A final rejection issued on all claims, except claims 55-58. Ex. 1017, 37-38. The patentee narrowed each originally issued independent claim to include the new limitation found in allowable claims 55-58 and canceled claims 42-81. Ex. 1018, 2-6, 9. Amended claims 1-41 were found patentable (Ex. 1019, ¶3). A reexamination certificate issued on May 21, 2014 (Ex. 1020).

### 3. Prior Proceedings Involving Related Patents

The ’923 patent is part of a family of patents that ultimately claim priority to U.S. Application No. 09/694,712. Related U.S. Patent Nos. 7,868,912 (“the ’912 patent”) (Ex. 1034) and 8,564,661 (“the ’661 patent”) (Ex. 1035) are in the family.

The ’912 patent was also challenged in an *inter partes* reexamination by Bosch. Ex. 1024. Like the ’923 patent Bosch reexamination, the parties settled before any action by the examiner. Ex. 1025. The ’912 patent was also challenged in an *ex parte* reexamination. Ex. 1026. The Patent Office instituted the reexamination on all claims on multiple grounds. Ex. 1027; Ex. 1028. After back and forth with the Patent Office (Ex. 1029; Ex. 1030; Ex. 1031), a reexamination certificate issued (Ex. 1032).

The ’661 patent was not challenged in any reexaminations; but it was challenged in two IPRs brought by Petitioners, IPR2018-00138 and IPR2018-00140 (“the Related IPRs”). The Related IPRs were filed on October 31, 2017, instituted on June 1, 2018, and are currently pending.

## C. Level of Skill

A POSITA would have (i) a Bachelor of Science degree in electrical engineering, computer engineering, or computer science, with approximately two years of experience or research related to video processing and/or surveillance systems, or (ii) equivalent training and work experience in computer engineering and video processing and/or surveillance systems. Ex. 1005, ¶76-80.

## V. CLAIM CONSTRUCTIONS

Claims are to be given their "broadest reasonable construction in light of the specification" ("BRI"). 37 C.F.R. § 42.200(b). Terms not discussed here should be given their ordinary and customary meaning as understood by a POSITA in light of the BRI standard and need not be further construed.

This Petition adopts the definitions provided in the Definitions section of the patent (Ex. 1001, 3:21-4:17) for any claim terms listed there, except as expressly explained below.

### A. "attributes of the object" (claims 1-7, 9-19, 22-28, 30-41); "attributes of each of the detected first and second objects" (claims 8, 29); "attributes of the detected object" (claims 20, 21)

The '923 patent describes "attributes" or "primitives" (which the patent uses interchangeably) as "observable" characteristics of an object. Ex. 1001, 7:6-7. Such characteristics include physical characteristics, *e.g.*, "a classification; a size; a shape; a color; a texture; a position." *Id.*, 7:8-12. Attributes can also represent actions or activities of the object, such as "a speed, an internal motion, a motion…" *Id.* Examples include: "appearance of an object, disappearance of an object, a vertical movement of an object, a horizontal movement of an object…." *Id.*, 7:37-40.

Accordingly, "attributes" as used in the claims should be construed as "characteristics associated with an object." Ex. 1005, ¶88-91.

**B. "new user rule" (claims 1-41)**

Each independent claim recites a "new user rule." Claim 1 recites "identifying an event of the object … by applying the ***new user rule*** to the plurality of detected attributes." The term is not defined in the claims. But dependent claims 2 and 23 provide that "selecting a new user rule" is "selecting a subset of the plurality of attributes for analysis."

The specification does not use the term "new user rule." But it does refer to "event discriminators." Ex. 1001, 4:63-64 ("an operator is provided with maximum flexibility in configuring the system by using event discriminators."); Ex. 1005, ¶93. Attributes/primitives are detected and archived, and then "event occurrences are extracted from the video primitives using event discriminators." Ex. 1001, 10:58-64; *see also* Fig. 4, 10:66-11:1. Event discriminators refers to one or more objects interacting with one or more spatial and/or temporal attributes. *Id.* at 7:1-12. Thus, the event discriminators in the patent perform the same function of allowing the user to identify events from attributes as the claimed "new user rule." Ex. 1005, ¶94.

Accordingly, "new user rule" should be construed to mean "a specified combination of a set of attributes for identifying an event." *Id.*, ¶95.

### C. Independence-based limitations (claims 1-41)

During the ’923 patent’s reexamination and in the Related IPRs, Avigilon argued that the claimed “independence-based” limitations distinguish its invention over the prior art. And Avigilon argued that the limitations have the following three requirements (1) identifying an event that refers to one or more objects engaged in an activity by analyzing the detected attributes; (2) the detected attributes are independent of the event identified; and (3) the identified event is not one of the detected attributes. IPR2018-00138, Paper No. 11, 25 (September 4, 2018); Ex. 1016, 37-39.

Petitioner addresses each of these concepts below using Avigilon’s numbering scheme. However, point (2) is addressed last for ease of analysis.

#### 1. Independence Argument (1)

Avigilon asserts that these limitations require identifying an event by analyzing the detected attributes. IPR2018-00138, Paper No. 11, 16-17 (September 4, 2018); Ex. 1016, 37-38. The claims of the ’923 patent articulate this concept rather generically, merely reciting that the event is identified “by ***applying*** the new user rule to the plurality of detected attributes.” Ex. 1001, Reexamination Certificate 1:45-47 (emphasis added). The ’923 patent does not use “apply/applying” in this context. *See, e.g., id.*, 4:64-5:1, 6:63-64, 7:2-6, 10:63-64, 10:66-7:1. Instead, the claims use the claim term in a purely functional sense according to its ordinary meaning. Thus, a POSITA would understand that the claimed "applying" would encompass any mechanism for analyzing the detected attributes to determine if they satisfy the user rule criteria, *e.g.*, querying a database. Ex. 1005, ¶¶100-103.

### 2. Independence Argument (3)

Argument (3) corresponds to claim language that the identified event is not one of the detected attributes. *See, e.g.*, Ex. 1001, Reexamination Certificate 1:44-45. This argument traces from Avigilon's efforts to distinguish the claims over the prior art (Ex. 1005, ¶¶106-107), and results in a departure from the basic definition of "event" in the "Definitions" section of the '923 patent (Ex. 1001, 3:44-46).

Specifically, Avigilon admitted that the '923 patent specification considers single activity attributes to be events by explaining in the reexamination that:

> the specification of the '923 patent discloses ***some identified events that are the same as a detected attribute***. *See* '707 application at ¶ 98 ("an object appears").

Ex. 1016, 38 (emphasis added). This example—"an object appears"—identifies any object that "appears." As such, the event "an object appears" merely identifies every occurrence of the "appear" attribute. Ex. 1005, ¶¶107-108.

But to distinguish the claims over prior art like *Courtney*, which was applied in the reexaminations, patentee argued that the claim language requires that the ***claimed*** event is more than a single attribute:

> the specification of the ’923 patent also discloses events that are not detected attributes. *See, e.g., id* at ¶ 98 (“a person appears; a red object moves faster than 10 m/s”); & ¶ 99 (“two objects come together; a person exits a vehicle; a red object moves next to a blue object”). The claims of the ’923 patent require identification of an event that is not a detected attribute and are silent regarding identification of an event that is a detected attribute. *See* Zeger Dec., ¶ 56.

Ex. 1016 at 39. Here, the patentee identified “a person appears” as an event that is within the scope of the claim because it is not merely a single event attribute. Instead, this event requires two attributes, the “appear” activity attribute plus a “person” object classification attribute. Ex. 1005, ¶109. Thus, patentee admits that single activity attributes are not events within the scope of the claim—although they would be events in the context of the patent disclosure—and two attribute events are events within the scope of the claim. *Id.*

In the Related IPRs, Avigilon has attempted to distinguish the prior art of this Petition by conflating single activity attributes with events by arguing that activities like appear, enter or exit, when recorded by the prior art are merely pre-determined events. IPR2018-00138, Paper No. 11, 7-10, 52-53 (September 4, 2018). That argument is completely inconsistent with the disclosure of the ’923 patent, which records those same activities. Ex. 1001, 3:31-33, 7:6-10, 7:37-41. Thus, the fact that a prior art system records activity attributes does not provide an adequate basis to distinguish that art from this limitation. Ex. 1005, ¶¶110-111.

Moreover, Avigilon's argument is inconsistent with the language of the '923 patent claims reciting "identifying an event of the object that is not one of the detected attributes of the object." This limitation clarifies that ***the claimed event specified by applying the user rule*** cannot be a single activity attribute. *Id.*, ¶112.

Thus, the '923 patent limitation associated with Argument (3) requires that the ***claimed*** user defined "event" comprises a minimum of two attributes. *Id.*, ¶113.

### 3. Independence Argument (2)

Argument (2) asserts that the detected attributes are "independent" of the event that is identified. As an initial matter, it is important to understand that the "event identified" is the event specified by the user rule and therefore the requirements of Argument (3) above must apply to this event and it cannot not merely be one of the detected attributes. Ex. 1005, ¶114.

The claim language corresponding to this argument requires:

- "the plurality of attributes that are detected are independent of which event is identified" (claims 1, 8, 22, 29)
- "the attributes to be detected are independent of the event to be detected" (claim 20)
- "for identifying the event independent of when the attributes are stored in memory" (claim 9, 30)

The key here is that the ***claimed*** independence is between the detected attributes and the event that is defined by the user rule and identified. *Id.*, ¶116. For example, claim 1 requires:

> **[1.3]** selecting a ***new user rule*** after detecting the plurality of attributes
>
> **[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 user rule*** to the plurality of detected attributes …
>
> **[1.5]** wherein the plurality of attributes that are detected are ***independent of which event is identified***,

As stated in [1.5], the claimed attributes are required to be independent of "which event is identified." *Id.*, ¶117. That identified event has antecedent basis in [1.4] where it is specified as the event defined by the claimed new user rule. This claim language should, therefore, be understood to require that the detection of attributes is independent from, *i.e.*, not affected by, the user rule that tasks the system. *Id.*; *see also* Ex. 1001, 6:64-67.

If a user rule can define an event that is an arbitrary mix of detected attributes and the definition of the event by the user rule is not used to alter the selection of attributes that are collected, the limitation is met. Ex. 1005, ¶118. Accordingly, the proper construction of this limitation merely requires that ***the event detection process does not alter the attribute detection process***. *Id.*

In the Related IPRs, Avigilon has argued that this limitation should be construed to mean "the plurality of detected attributes are detected ***without regard to or knowledge*** of ***a predefined/predetermined list of events of interest***" amongst which at least one event is identified. IPR2018-00138, Paper No. 11, 29 (September 4, 2018) (emphasis added). This proposed construction is not supported by the claim language, specification or prosecution history of the ’923 patent. *See also* Ex. 1038, ¶¶ 50, 56, 60 (expert identified no support in the intrinsic record). The claim only requires independence from the event identified by the user rule, not a list, and the claim construction should not add this unstated concept. Ex. 1005, ¶119.

Further, the ’923 patent specification disclosure merely states that in the disclosed embodiment tasking is optional and that without tasking the system would still detect attributes. *See* Ex. 1001, 6:61-67; IPR2018-00138, Paper No. 7, 14-15 (March 2, 2018). This passage does not disclose what it means to have “regard for” the events and who or what is prohibited from having “knowledge” of events of interest. Ex. 1005, ¶¶121-122. Neither Avigilon’s construction nor the patent specification answers these questions and that construction should be rejected on that basis. *See* IPR2018-00138, Paper No. 11, 24-29 (September 4, 2018).

Indeed, the ’923 patent’s claimed real-time detection feature contradicts Avigilon’s argument because, once a new user rule is defined, the system operating in a real-time mode has knowledge of that ultimate event to be determined, and all of this occurs ***before*** the attributes are even detected. Ex. 1005, ¶123.

Avigilon further sows confusion as to what constitutes an "event" in the claimed system by arguing that prior art, like *Courtney*, is distinguishable over the claims because it detects ***predefined*** "events." IPR2018-00138, Paper No. 11, 11-12 (September 4, 2018). The problem with the Avigilon's argument is that it does not apply the definition of the event specified in the claims according to the Argument (3) claim language, *i.e.*, that the claimed event is not a single attribute. Ex. 1005, ¶124.

Under that definition, what *Courtney* calls an event is the same as a single activity attribute in the '923 patent, *i.e.*, not really an event under Avigilon's claims. *Id.*, ¶125. To overcome rejections in the reexamination, patentee argued that *Courtney* was distinguishable because when it analyzed video it merely detected "events." Ex. 1016 at 49. While it is true that *Courtney* does use the word "events," it refers to ***single activity attributes***, such as appear, disappear, enter, and exit. Ex. 1021, 10:52-61. The '923 patent detects these same things but happens to label them activity ***attributes***, as opposed to events. Ex. 1001, 3:30-33. And the patentee clarified that single activity attributes should not be interpreted as events in the context of its claims. Ex. 1016, 38-39. Thus, the mere fact that *Courtney* happened to label certain detected activities like "appear" as an "event" (Ex. 1021, 10:52-61) rather than an "attribute" is nothing more than a semantic difference. It is not a patentably significant technical distinction compared to the ’923 patent claims. Ex. 1005, ¶126.

To distinguish *Courtney*, the patentee argued that referencing events by location or time does not meet the “identifying events” requirement in the claim. Ex. 1016, 49. The patentee argued that *Courtney* merely queries its single activity attribute “events” and adds a time and/or location attribute, *e.g.*, an object appears at a certain time and/or location. *Id.* This argument is wrong. Ex. 1005, ¶¶127-129.

First, there is ***no*** claim limitation that expressly prohibits defining an event by adding location and/or time attributes to a single activity attribute. The mere use of the word “independent” certainly does not dictate this concept. Rather, the claims require the event definition includes temporal attributes. Ex. 1001, Reexamination Certificate 1:37-39; *see also id.* 16:22-25. Second, nor does patentee’s construction support this concept. That construction simply requires the system to detect attributes without knowledge of predefined “events,” and ***not*** predefined “activity attributes.” This proposed construction does not preclude adding location and/or time attributes to a single activity attribute to define an event. Third, patentee’s argument is contrary to the patent’s express disclosure of events that are a collection of an object’s single activity attribute and location and/or time. *See*, *e.g.*, Ex. 1001, 9:2 ("an object ***appears*** at ***10:00 p.m.***") (emphasis added), 5:1-5.

If Avigilon maintains this time or location argument, it should at most only exclude the claim from covering a system that can ***only*** identify an "event" as a single predefined activity attribute, plus a time attribute, and/or a location attribute. This, however, would not distinguish over the prior art presented in this Petition, which is capable of much more sophisticated *ad hoc* event definitions. Ex. 1005, ¶130.

The proper construction of this limitation merely requires that the ***event detection process does not alter the attribute detection process***. *Id.*, ¶131.

### D. "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" (claims 1-19, 22-29); "wherein the analysis of the combination of the attributes to detect the event comprises analyzing only the combination of the attributes" (claims 20-21); "wherein the applying the selected new user rule to the plurality of attributes stored in memory comprises applying the selected new user rule to only the plurality of attributes stored in memory" (claims 30-41)

In the '923 reexamination, the patentee added the limitation "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" to distinguish the *Day-I* and *Day-II* references. Ex. 1018, 3-6, 9. Patentee argued that this language distinguished the *Day* references because "the queries of *Day-I* are not applied to the attributes stored in the VSDG alone but are applied to object-oriented abstractions." Ex. 1016, 78-79. Patentee further argued that *Day-I* stated that "all these queries generally require processing of various combination [sic] of ***object hierarchy*** (shown in Figure 5)." *Id.* (emphasis added). Patentee made similar arguments as to *Day-II*. *Id.*

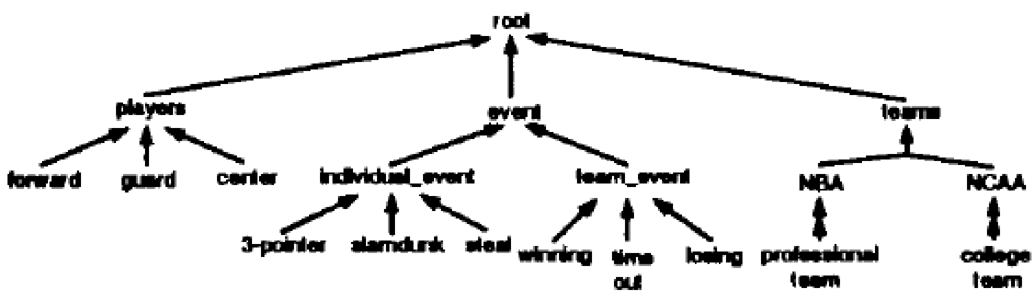



**Figure 5: Fan's view**

Patentee's arguments to overcome the *Day* references are not well founded. Ex. 1005, ¶133. As discussed in Section IV(B)(2), patentee's citation to the patent does not support its argument that the patent describes only using primitives and not abstractions. Indeed, the patent expressly discloses that the system can process abstractions, but this was never cited to. Ex. 1001, 8:16-17, 50-53. There is no disclosure explaining what using primitives versus abstractions means or how one would embody a working system that did not process abstractions, as the patentee argued to overcome the *Day* references. Ex. 1005, ¶¶133-134.

Nevertheless, the patentee essentially argued that the *Day* references were distinguishable because they ***always*** require the processing of abstractions of the attributes and not the attributes themselves when applying a user rule. *Id.*, ¶¶135-136. This is because *Day* always references an object hierarchy structure, *e.g.*, a tree structure, in performing a search. *See* Ex. 1022, Fig. 5.

The claim language requires that the applying step ***comprises*** applying the new user rule to only the detected attributes. Ex. 1005, ¶137. The use of the open-ended term "comprising" encompasses systems that employ searches of object-oriented abstractions, so long as they can also employ searches of only the attributes themselves.

In sum, this limitation should at most only limit claims as excluding 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. *Id.*, ¶138.

### E. Means-plus-function elements (claims 9-19, 30-41)

The following elements should be construed under pre-AIA 35 U.S.C. § 112(6) because they recite "means for" without reciting sufficient structure to perform the recited functions.

#### 1. "means for detecting an object in a video from a single camera" (claim 9); "means for detecting first and second objects in a video from a single camera" (claim 30)

The '923 patent recites two corresponding structures for these limitations, both of which are conventional motion and/or change detection algorithms that are utilized on a computer system or equivalent video processing system to detect objects: 1) Ex. 1001, 5:61-64, 9:33-35 (Fig. 5 (51)); and 2) 5:61-64, 9:39-41 (Fig. 5 (52)). Ex. 1005, ¶140.

#### 2. "means for detecting a plurality of attributes of the object …" (claim 9); "means for detecting a plurality of attributes of the object…" (claim 30)

The '923 patent recites a corresponding structure for these limitations, which is a conventional computer-vision algorithm that is utilized on a computer system or equivalent video processing system for detecting attributes: Ex. 1001, 5:61-64, 10:49-51 (Fig. 5 (57)). Ex. 1005, ¶141. Examples of a physical attribute of an object include size, shape, color and texture, etc. Ex. 1001, 7:8-9. Examples of a temporal attribute of an object include "every 15 minutes," "between 9:00 p.m. to 6:30 a.m.," "less than 5 minutes," etc. *Id.*, 8:32-36.

#### 3. "means for selecting a new user rule after the plurality of detected attributes are stored in memory" (claim 9)

The '923 patent recites a corresponding structure for this limitation, which is a conventional user interface software and "I/O devices" such as "a keyboard; a mouse; a stylus; a monitor," utilized on a computer system or equivalent video processing system: Ex. 1001, 5:61-64; 6:23-28, 6:61-64 (Fig. 2 (23)), 15:11 (Fig. 9 (91)). Ex. 1005, ¶142.

#### 4. "means for identifying an event…" (claim 9); "means for identifying an event of the first object interacting with the second object…" (claim 30)

The '923 patent recites a corresponding structure for these limitations, which is a conventional query mechanism utilized on a computer system or equivalent video processing system to detect an event: Ex. 1001, 5:61-64, 10:63-64 (Fig. 1(44)), 10:66-11:1, 9:14-17, 14:57-60, 15:7-10, 14:63-66, 8:65-67. Ex. 1005, ¶143. Any functional limitations containing the "new user rule" limitation or the independence-based limitations should be construed according to the claim constructions set forth in Sections V(B) and V(C), respectively. A POSITA would readily understand that the other functional limitations are performed by the query mechanism and no further construction is necessary. *Id.*

## VI. OVERVIEW OF PRIOR ART

### A. Kellogg

*Kellogg* discloses a visual memory system that tracks objects in a video, detects and stores information about the objects, and responds to user queries specifying events concerning those objects. Ex. 1003, Abstract, 69. The system includes an "image processing" to detect objects and attributes, a "visual memory" where attributes are stored, and a "graphical query interface" where a user defines query specifications for identifying events based on the attributes. *Id.*, Figure 4-1. *Kellogg*'s system "provides a powerful and expressive [querying] mechanism for retrieving information" that is "designed to meet a wide variety of retrieval needs, providing flexibility in specifying objects of interest." *Id.*, 53.

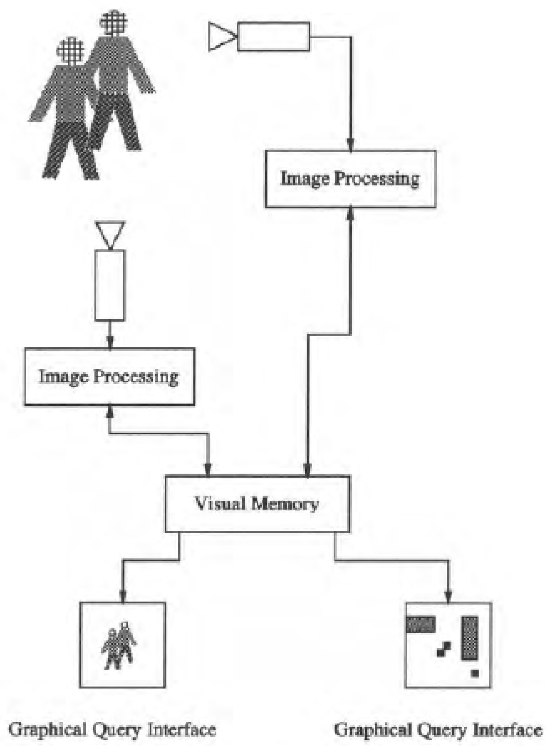



Figure 4-1: Scene monitoring prototype

*Kellogg's* system detects and stores object attributes. The system detects an object's area, duration, trajectory (*id.*, 22); the object's class, centroid, orientation, bounding box (*id.*, 24); relative special attributes ("west" or "near") (*id.*, 30); time stamps or intervals for the valid times when the object existed (*id.*, 36-37, 52); and height of a person (*id.*, 71).

*Kellogg* uses database queries to detect events. *Id.*, 53. Rules "are implemented as part of the query language to allow the query language to optimize object retrieval." *Id.*, 54. *Kellogg*'s query mechanisms provide "great flexibility in spatial and temporal query specification," allowing queries to "include[] spatial or temporal keywords," "spatial or temporal object[s]," or even "the result[s] of another query." *Id.*

*Kellogg* discloses querying events that are not merely a single detected attribute. Users can define events of their choosing, such as determining when a person (an object classification) does or does not intersect an arbitrary user-drawn rectangle ("enter area activity") by comparing the persons' position. *Id.*, 54-58, *see also,* 58 (finding when objects are in the scene during the same time). More complex spatiotemporal searches are also disclosed, which allow, *e.g.*, identifying an "Approach Event" that finds "all objects that came within 3 units of a given object on its trajectory during a certain set of valid times." *Id.*, 63. Other examples of *ad hoc* user rules disclosed by *Kellogg* include, "[w]atch for anything that comes within 3 feet of that button" (*id.*, 68), or determine if "anybody [came] into the room between 12:00 and 1:00" (*id.*, 80).

*Kellogg*'s user can "specify that an alarm should fire only if an object remains in a region for a suspicious amount of time," which is the same "loiter" event detected by the '923 patent. *Id.*, 80; *see* Ex. 1001, 5:1-5.

## B. Brill

*Brill* discloses an Autonomous Video System ("AVS") system for moving object detection and event recognition. Ex. 1004, 4. *Brill*'s AVS system is shown in Figure 1:

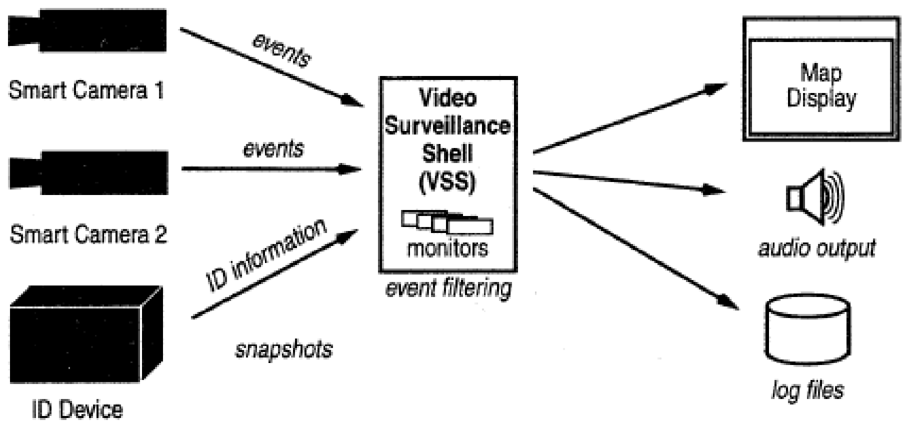



Figure 1: AVS system diagram

As illustrated in Figure 1, smart cameras process and analyze video to send activity attributes and other attributes (*e.g.*, object ID (Ex. 1004, 11), object type, location, and time stamp (*id.*, 13)) to the Video Surveillance Shell ("VSS"). At the VSS, user-entered "monitors" are applied to detect events according to the '923 patent. *Id.*, 13.

*Brill* teaches a graphical user interface that allows a user to define and submit queries to identify events. *Id.*, 13, Figure 11.


Figure 11: Selecting a type of simple event

In Figure 11, the user defines a rule for a new “Loiter by the door” event, which involves a “person” object engaging in a “loiter” activity. Ex. 1005, ¶161. It further requires an “outside the door” region and a duration of “5.0” seconds. *Id.* For this specific rule, the “days of week” and “time of day” attributes are not analyzed, but a user can include those attributes to narrow down the search. *Id.*

### C. Motivation to Combine

A POSITA would have found it obvious to combine *Kellogg’s* system with the features of *Brill*’s system. Ex. 1005, ¶167. Indeed, the Board has found a motivation in the Related IPRs. IPR2018-00138, Paper No. 8, 10, 18-19 (June 1, 2018).

As demonstrated by the hundreds of references cited on the face of the ’923 patent, the state of the art was quite crowded. Ex. 1001, References Cited; *see also* Ex. 1005, ¶167. A POSITA would be aware of object detection methods, attribute detection methods, and querying mechanisms, like those disclosed in *Kellogg* and *Brill*. Ex. 1005, ¶167. A POSITA would have combined elements of *Kellogg* and *Brill* to provide enhancements or achieve design objectives, while yielding predictable results. *Id.*

*Kellogg* teaches detecting one or more objects in a video obtained from a single camera. Ex. 1003, 30-31, 77, 79, Figure 3-5, 4-9. Those objects could be multiple people or a person and vehicle in a single field of view. *Id.*, 56-57, 79. *Kellogg* also contemplates detecting interactions between those objects. *Id.*, 65-67; *see also* Section VIII(A)(5)(f).

*Brill* teaches, among other things, an enhanced event detection platform that reliably handles recognizing interactions of multiple objects, especially human-vehicle and human-human interactions. Ex. 1004, 6; Ex. 1005, ¶169.

*Brill* teaches improvements to object tracking so that tracking is not lost when a person's image overlaps another object's, *e.g.*, a car. Ex. 1004, 6-9; Ex. 1005, ¶170. This would cause the object to appear to merge, resulting in the loss of tracking until the person walks away from the other object. Ex. 1005, ¶170. A POSITA who employed *Kellogg*'s system to monitor a human and a vehicle, and any interaction thereof, would have been aware of this issue. *Id.*; Ex. 1003, 79 (the alarm region monitors both human and vehicle).

*Brill* teaches that its "new approach involve[s] 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 motivated to combine the teachings of *Brill* with *Kellogg*'s monitoring system to solve the loss of tracking issue. Ex. 1005, ¶170.

A POSITA using *Kellogg*'s system to monitor human-human interaction also faced a separate issue. *Id.*, ¶171. As described in *Brill*, a POSITA knew that it is difficult to monitor movements of one or more people in a scene because they move unpredictably, move close to one another, and occlude each other. Ex. 1004, 14. Indeed, *Kellogg* acknowledges these types of uncertainties. Ex. 1003, 31-35. As *Brill* explains, when two people are in a single scene, it was difficult to maintain the separate tracks of the two people once they merge into a single large region. Ex. 1004, 14. A POSITA employing *Kellogg*'s system would have faced this issue. Ex. 1005, ¶171.

*Brill* introduces a new method which maintains an estimate of the size and location of the objects, and that creates a separate image which approximates the probability that the object intersects that pixel location. *Id.*, ¶172; Ex. 1004, 14-15, Figure 15.

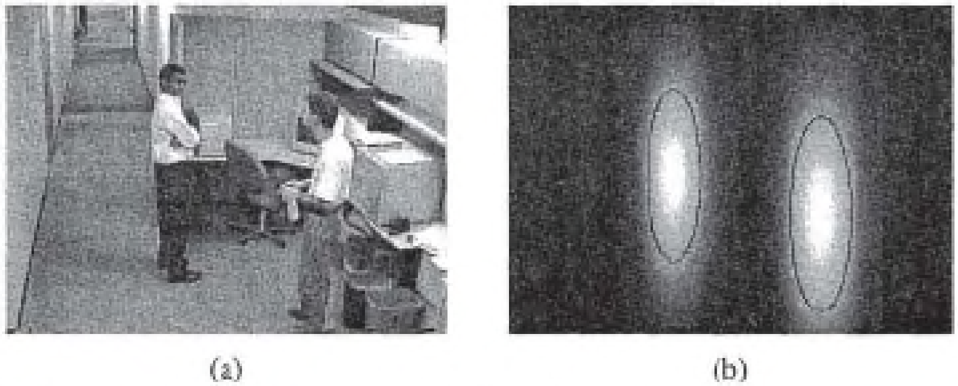


Figure 15: Probability image for the locations of the people in the scene

*Brill* teaches that even after an occlusion, the objects are reliably detected and tracked by relying on non-overlapping areas. Ex. 1004, 15. A POSITA would have been further 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. Ex. 1005, ¶173.

The motivation to combine *Kellogg* and *Brill* is further evidenced by the fact that the AVS systems were developed by Texas Instruments. *Id.*, ¶174. Indeed, the art confirms that the visual memory database of *Kellogg* was combined with *Brill's* AVS system. *Id.*; *see also* Ex. 1007, ¶¶40-44.

Both references provide express teaching, suggestion, and motivation to combine *Brill*'s teachings with *Kellogg*'s system. Ex. 1005, ¶175. Moreover, as both relate to the Texas Instruments AVS systems, a POSITA would be able to predictably combine the teachings of the two references, without requiring extensive modification to the overall system. *Id.* *Brill* would be an obvious source because, like *Kellogg*, *Brill* is directed to video surveillance systems using object detection methods, attribute detection methods, and querying mechanisms. *See, e.g.*, Ex. 1003, 24-25, 50, 54, 62-63, 68-69, 71, 77; Ex. 1004, 6-9, 12-14. Similar to *Kellogg's*, *Brill's* system can detect multiple objects and their interactions in a single video scene. *See, e.g.*, Ex. 1003, 62-63; Ex. 1004, 12-13. Given the similarities of their subject matter and teachings, a POSITA would have immediately recognized *Brill*'s advanced multi-object detection and tracking method would have readily worked in *Kellogg*'s system with a reasonable expectation of success. Ex. 1005, ¶175.

## VII. THE PETITION PRESENTS NEW ISSUES OF PATENTABILITY AND THE BOARD SHOULD NOT EXERCISE ITS DISCRETION TO REJECT THE PETITION

None of the prior art references used in this Petition were cited in a ground for rejection during prosecution, or any reexamination of the '923 patent. The Board has already found that the prior art presented herein raises new patentability issues and is not substantially the same as the prior art references previously considered by the PTO against this patent family. IPR2018-00138, Paper No. 8, 6-7 (June 1, 2018).

### 1. Dispute Regarding the Independence-Based Claim Elements in the Prior Proceedings

In the *ex parte* reexaminations of the '923 and '912 patents, the patentee argued that the primary references failed to disclose the "independence-based claim elements." Ex. 1016, 46-52, 54-61; Ex. 1029, 40-46, 51-54. These arguments were directed to the independence claim construction issues discussed above in Section V(C), including the attempt to distinguish the '923 patent from *Courtney*.

In the Related IPRs, the crux of Avigilon's Preliminary Response argument was that none of the cited references—namely *Kellogg*, *Brill* and *Dimitrova*[2]—disclose the "independence-based claim elements" because they are no different than *Courtney*. IPR2018-00138, Paper 7, 19-28 (March 2, 2018); *see also* IPR2018-00140, Paper 7, 19-22 (March 2, 2018). The Board rejected that argument. IPR2018-00138, Paper No. 8, 6-7 (June 1, 2018). In Avigilon's Response, Avigilon argued that *Kellogg* is like *Courtney because* it merely indexes predefined events. IPR2018-00138, Paper No. 11, 36-37 (September 4, 2018). Avigilon's argument is wrong.

Unlike *Courtney*, *Kellogg* does ***not*** apply its queries to a predefined list of events. Ex. 1005, ¶187. *Kellogg* only stores basic attributes, such as object's area, duration, and trajectory (Ex. 1003, 22), object's class, centroid, orientation, and bounding box (*id.*, 24), relative spatial attributes (*id.*, 30), time stamps or intervals

[2] *Dimitrova* is the primary reference relied upon by the Petitioner in IPR2018-00140.

for the valid times when the object existed (*id.*, 36-37, 52), and volume or height (*id.*, 25, 71). With these basic attributes, *Kellogg* allows a user to later select any arbitrary collection of attributes to define a new event. Ex. 1005, ¶187. For example, after video attributes are stored, a user might create a completely new event to uncover, such as whether an object classified as a person intersected with a new, arbitrarily defined rectangular region. Ex. 1003, 55 ("[s]elect p from Person where p intersects %rectangle."). Notably, *Kellogg's* system does not and could not predefine or index this intersection activity because the user had not come up with it yet. Ex. 1005, ¶187. Instead of being predefined, this intersection is computed based on more fundamental movement attributes. *Id.*

Moreover, unlike *Courtney's* query, *Kellogg's* query does ***not*** limit the search to a single activity attribute, but also searches for an object classification attribute, *e.g.*, person. Ex. 1003, 55 ("[s]elect p from Person where p intersects %rectangle."). This is the same additional attribute that Avigilon identified as characterizing its invention. *See* Ex. 1016 at 38-39 ("the specification of the '923 patent also discloses events that are not detected attributes … [*e.g.*,] 'a ***person*** appears'") (emphasis added); Ex. 1005, ¶188.

In contrast, *Courtney's* query cannot identify an *ad hoc* after-the-fact intersection event like this because it is not one of *Courtney's* predefined events (appearance, disappearance, entrance, exit, deposit, removal, motion and rest). Ex.

1021, 10:50-61; Ex. 1005, ¶189. As the Board has already found, *Kellogg* teaches "independence" and this Petition presents substantially different art than what was before the PTO during the reexamination.

As explained in Section VIII(A)(1)(d), *Kellogg* discloses embodiments that apply a new user rule "only" to the detected attributes themselves without using any higher-level abstractions. Indeed, *Kellogg* discloses a specific embodiment where the attribute data is stored in a "bucket index," which is merely a collection of object attributes without any higher-level organizational structure, which is fundamentally different than the hierarchy of *Day-I* or *Day-II*. *See* Section V(D); Ex. 1005, ¶190.

As it did in the Related IPRs, the Board should institute this IPR because *Kellogg* and *Brill* provide more than a reasonable likelihood that the challenged claims are invalid. IPR2018-00138, Paper No. 8 (June 1, 2018).

# VIII. CLAIMS 1-41 OF THE '923 PATENT ARE UNPATENTABLE OVER THE PRIOR ART

## A. *Kellogg* Anticipates Claims 1-41

### 1. Independent Claim 1

*Kellogg* discloses the method of claim 1.

**a. “detecting an object in a video from a single camera”**

*Kellogg* discloses detecting objects, tracking them, and storing information about them in the visual memory. Ex. 1003, 68-69. When a new object is detected in a video, an object identifier (“OID”) is assigned to that object. *Id.*, 50.

*Kellogg* discloses “real-time processing of CCD camera images” and software that “tracks ***people*** walking in its ***field of view***.” *Id.*, 77 (emphasis added). The plural word “people” emphasizes that more than one object is detected from a video from a single camera. Ex. 1005, ¶198. *Kellogg* also contemplates detecting at least two different objects from a video from a single camera. *Id.*; *see, e.g.*, Ex. 1003, 79 (monitoring both people and vehicles at an outdoor alarm region), Figure 4-9 (monitoring three objects—each represented as small squares—from a single camera (red)).

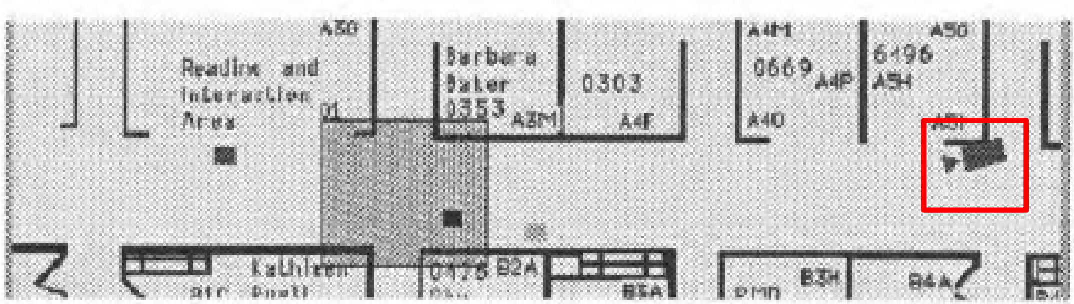



Figure 4-9: Graphical query results

As shown in Figure 3-5, *Kellogg* detects the spatial relationship among different objects, *i.e.*, first and second objects, and keeps track of them in a single scene. Ex. 1003, 30-31; *see also* 21 (index tracks “centroids of moving objects”).

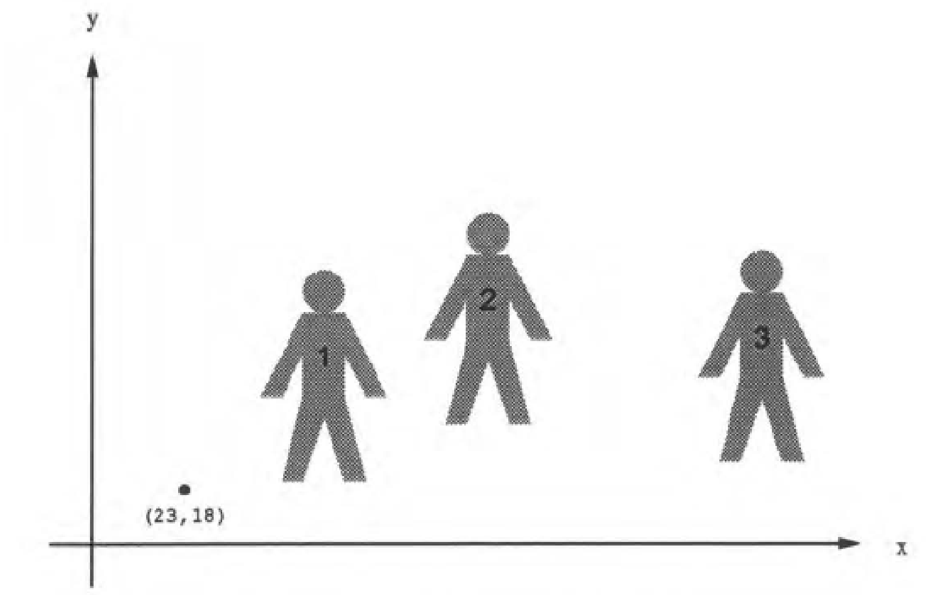



Figure 3-5: Relative spatial objects

Accordingly, a POSITA would understand that *Kellogg* discloses this element. Ex. 1005, ¶¶196-201; *see also* Ex. 1001, 9:35-38, 9:44-48 (admitting many object detection methods were known).

**b. "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"**

*Kellogg* discloses detecting a plurality of attributes of the detected object by analyzing the video. Ex. 1005, ¶205. For example, the "software estimates the positions and heights of people" from the field of view. Ex. 1003, 77.

*Kellogg* teaches detecting several physical attributes, including an object's classification (*e.g.*, person, cube), centroid (position of the center of the object), orientation (position of the object), and bounding box (size, shape, and position). Ex. 1003, 24, 63; Ex. 1005, ¶¶202-205. Additional physical attributes can be detected for certain types of objects, such as the volume of a cube or the height of a person. Ex. 1003, 25, 71.

*Kellogg* discloses the temporal attributes "*TemporalInterval,*" representing the interval time(s) for which an object is present, "*VMTime*," representing the single time that an object is present (*i.e.*, a timestamp), and "*RelativeTemporalObject,*" representing the existence of an object in relation to other objects may be detected. *Id.*, 36-41; Ex. 1005, ¶¶206-207.

These attributes represent characteristics of the detected object. *See also*, Section V(A).

### c. "selecting a new user rule after detecting the plurality of attributes"

*Kellogg* discloses a system tasking mechanism to allow users to create *ad hoc* queries using a set of attributes, *i.e.*, selecting a new user rule. Ex. 1005, ¶¶209-210, 215-216. *Kellogg* provides many examples of creating user-defined queries ***after*** detecting the plurality of attributes. *Id.*

One example is the query (below) that "tracks all objects that ***came*** within 3 units of a given object on its trajectory during a certain set of valid times" (hereinafter the "Approach Query"). Ex. 1003, 62-63 (emphasis added). Use of the past tense "came" further emphasizes that a user creates a new rule seeking to extract a new event of interest from previously detected attributes. Ex. 1005, ¶211.

```
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)
```

The Approach Query specifies a combination of a set of physical and temporal attributes together with object classification attributes for identifying an "Approach Event" between person p and any person q by searching for any person q that came within a certain distance of person p at a given time period. *Id.*, ¶212. The query searches the centroid of each object, the centroid trajectory of each object, and the distance between the moving centroids. *Id.* It searches for trajectories within certain periods, *e.g.*, time-1 and time-2. *Id.* It also searches for the object classification attribute, *i.e.*, person, for both objects. *Id.* These attributes are detected before a user creates the Approach Query. *Id.*

If Avigilon argues that the term "new user rule" requires setting a response, with which the Petitioner disagrees, *Kellogg's* system provides this feature by allowing a user to enter a "delay specification that indicates how long an object must remain in that region before the system triggers an alarm." Ex. 1003, 79-80; Ex. 1005, ¶¶213-216.

**d. "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 user rule to the plurality of detected attributes, 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"**

The limitation "identifying an event of the object that is not one of the detected attributes of the object" means "the user defined event comprises a minimum of two attributes." *See* Section V(C)(2). The Approach Query identifies an Approach Event between the person p and any person q. Ex. 1003, 62-63. The query applies at least two different attributes from the plurality of detected attributes, *e.g.*, centroid trajectories of objects p and q, object classifications (person) of objects p and q, time-1, and time-2, etc. Ex. 1005, ¶218. The query is applied ***after*** the plurality of attributes are detected and ***after*** creating the query because *Kellogg's* system was designed to record basic attributes so that a search can later be created based on any arbitrary subset of the recorded attributes. *Id.*, ¶220; Ex. 1003, 53-54, 78.

The limitation "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" requires that the prior art have the ability to search ***only*** the attributes themselves. *See* Section V(D); Ex. 1005, ¶221. *Kellogg* can search only the attributes themselves and does not require traversing a tree structure of abstractions to search the detected attributes. *Id.*, ¶222. *Kellogg* discloses that the attribute data can be stored in a "bucket index." Ex. 1003, 83. A "bucket index simply maintains a list of all the objects stored in the visual memory … [and] answers a query by retrieving all the objects in its list and checking them against the query specification," demonstrating that the system can search only the attributes themselves. *Id.*; Ex. 1005, ¶222.

#### e. "wherein the plurality of attributes that are detected are independent of which event is identified"

This limitation requires that attribute detection is not impacted or affected by the event detection process. *See* Section V(C)(3). *Kellogg*'s Approach Query meets this limitation because it identifies the Approach Event by using multiple ***pre-collected*** activity, physical and temporal attributes. Ex. 1003, 62-63. For example, the query searches the pre-collected centroid trajectories of the two moving objects, and calculates the distance between the moving centroids. Ex. 1005, ¶223. These moving trajectories must be within certain time periods, *e.g.*, time-1 and time-2, which are also pre-collected. *Id.* The query also applies the object classification attribute, *i.e.*, person, for both objects. *Id.*

```
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)
```

None of the attribute detection process is affected by the identification of this Approach Event. *Id.*, ¶224. For example, the system collects the basic centroid attributes of the two objects regardless of the distance parameter set after the fact by the user. *Id.* Moreover, a user can adjust the distance parameter *ad hoc*, thus specifying what an Approach Event is, but none of the detected attributes will be affected. *Id.*

Other detected attributes can be mixed with these detected attributes to identify a completely different event. *Id.*, ¶225. The centroid attributes can be used in identifying events other than the Approach Event, such as identifying objects whose centroids are located in an arbitrary space using the "Intersects Query." *Id.*; Ex. 1003, 55.

Plus, *Kellogg* discloses that the queries are forensic in nature, where "[a] large amount of [stored] information could be established ***prior to*** application execution." Ex. 1003, 9 (emphasis added); *see also, id.,* 19 (visual memory updated without regard to the application that may later use the data, which "do not need to know how [the database] achieves its results").

Accordingly, *Kellogg's* attributes are detected and stored in the visual memory database without being affected by the event detection process. Ex. 1005, ¶227.

#### f. "wherein the step of identifying the event of the object identifies the event without reprocessing the video"

Because *Kellogg's* system has already stored the object attributes in its database, it needs only to search the stored attributes, without reprocessing the video, to find objects engaged in activities that meet the query specification. Ex. 1003, 53-54, 77, 79; Ex. 1005, ¶¶228-229.

#### g. "wherein the event of the object refers to the object engaged in an activity"

*Kellogg's* Approach Query identifies an Approach Event of person p with any person q where p and/or q are engaged in an action of moving along a certain trajectory. Ex. 1003, 62-63; Ex. 1005, ¶230. An activity "refers to one or more composites of actions of one or more objects." Ex. 1001, 3:31-33. Accordingly, the Approach Event refers to an object engaged in an activity. Ex. 1005, ¶230.

### 2. Claims 2, 4, 7[2][3], 13, 16, 23, 25, 28[2], 34, and 38

*Kellogg* discloses that the queries do not require analysis of all detected attributes, as recited in claims 2, 4, 7[2], 13, 16, 23, 25, 28[2],34, and 38. Ex. 1005, ¶231.

- "selecting the new user rule comprises selecting a subset of the plurality of attributes for analysis" (claim 2, claim 23)
- "no analysis is performed on at least some of the detected attributes to detect an event" (claim 4)
- "analyzing only a subset of the attributes stored in the memory" (claim 7[2], claim 28[2])
- "analyzing, of the plurality of attributes, only a selected subset of the plurality of attributes" (claim 13, claim 34)
- "analyzing a selection of individual ones of the detected plural attributes" (claim 16, claim 38)
- "do not cause the computer system to perform an analysis on at least some of the detected attributes to detect an event" (claim 25)

*Kellogg*'s system detects people's heights as an attribute. Ex. 1003, 77. But the Approach Query does not search for the heights of person p or any person q. *Id.*, 62-63. Accordingly, *Kellogg's* queries do not require analyzing all detected attributes, and disclose the above limitations. Ex. 1005, ¶232.

---

[3] [2] refers to the second element of the claim.

```
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)
```

### 3. Claims 3, 7[1][4], 17, 24, 28[1], and 39

The features added by claims 3, 7[1] 17, 24, 28[1], and 39 require that the detected attributes are stored in the system, and claims 17 and 30 further require that this storage happens prior to selecting a set of attributes for searching. Ex. 1005, ¶233.

- "the plurality of attributes that are detected are defined in a device prior to a selection of a subset of the plurality of attributes" (claim 3, claim 24)
- "storing the detected plurality of attributes in memory" (claim 7[1], claim 28[1])
- "the plural attributes detected by the means for detecting are defined in the video device independent of a selection of the detected plural attributes" (claim 17, claim 39)

*Kellogg's* "computer vision algorithms for a security system could analyze data provided by various cameras and ***store*** information in the visual ***memory***." Ex. 1003, 10 (emphasis added). "Applications could ***then retrieve*** this data to track objects, watch for suspicious events, and ***respond to user queries***." *Id.*

[4] [1] refers to the first element of the claim.

(emphasis added); *see* Ex. 1005, ¶234. *Kellogg* also discloses that the "query mechanism works on two levels, on disk [*i.e.*, longer-term storage] and in [working] memory." Ex. 1003, 54; *see also* 68. Either way, attributes are stored in memory, which is in a video device. Ex. 1005, ¶234. Accordingly, *Kellogg* discloses the above limitations. *Id.*

#### 4. Claims 5, 6, 15, 21, 26, 27, and 37

The additional features of claims 5, 6, 15, 21, 26, 27, and 37 require that the new user rule is applied to plural physical and/or temporal attributes. Ex. 1005, ¶235.

- "plurality of attributes include plural physical attributes and the method comprises applying the new user rule to a plural number of physical attributes" (claim 5, claim 26[5])
- "plurality of attributes include plural temporal attributes and the method comprises applying the new user rule to a plural number of temporal attributes" (claim 6, claim 27[6])
- "analyzing at least two selected physical attributes of the plurality of attributes" (claim 15, claim 37[7])

[5] No meaningful difference between claims 5 and 26. Ex. 1005, FN6.

[6] No meaningful difference between claims 6 and 27. *Id.*, FN7.

[7] Claim 37 recites the limitation in means-plus-function. The corresponding structure is the same structure identified in Section VIII(A)(15)(e), which performs the additional function. *Id.*, FN8.

- "a video device … which detects plural physical attributes and plural temporal attributes of the detected object upon analyzing the video; and then, selecting the new user rule to provide an analysis of a combination of the plural physical attributes and the plural temporal attributes to detect the event" (claim 21)

A plural number of physical attributes of an object are searched in the Approach Query, including at least: (1) the classification "person" of the two objects (2) the centroid of the two objects, (3) the centroids' trajectories, and (4) the distance between the moving centroids. Ex. 1003, 62-63. A centroid (which is essentially the position of an object) and object classification are identified as physical attributes in the '923 patent. Ex. 1001, 7:8-16; Ex. 1005, ¶¶236-238.

```
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)
```

*Kellogg*'s Approach Query applies a user rule to a plural number of temporal attributes as demonstrated by the two ***different*** temporal attributes, times-1 and times-2. Ex. 1003, 62-63; Ex. 1005, ¶239.

Accordingly, the Approach Query also analyzes plural physical and plural temporal attributes. Ex. 1005, ¶241.

### 5. Independent Claim 8

*Kellogg* discloses the method of claim 8.

**a. "detecting first and second objects in a video from a single camera"**

As explained in Section VIII(A)(1)(a), *Kellogg* discloses this element. Ex. 1005, ¶¶243-246.

**b. "detecting a plurality of attributes of each of the detected first and second objects by analyzing the video from said single camera, each attribute representing a characteristic of the respective detected object"**

*Kellogg* discloses detecting a plurality of attributes of each detected object by analyzing the video. *See* Section VIII(A)(1)(b); Ex. 1005, ¶247. Moreover, the "software estimates the positions and heights of ***people***" from the field of view. Ex. 1003, 77 (emphasis added to the plural of person). The plural attributes correspond to characteristics of the detected objects. *See* Section V(A).

**c. "selecting a new user rule"**

As explained in Section VIII(A)(1)(c), *Kellogg* discloses this element. Ex. 1005, ¶248.

**d. "after detecting the plurality of attributes, identifying an event that is not one of the detected attributes of the first and second objects by applying the new user rule to the plurality of detected attributes, 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"**

As explained in Section VIII(A)(1)(d), *Kellogg* discloses this element. Ex. 1005, ¶¶249-250.

**e. "wherein the plurality of attributes that are detected are independent of which event is identified"**

As explained in Section VIII(A)(1)(e), *Kellogg* discloses this element. Ex. 1005, ¶251.

**f. "wherein the step of identifying an event of the object comprises identifying a first event of the first object interacting with the second object by analyzing the detected attributes of the first and second objects, the first event not being one of the detected attributes"**

As discussed in Section VIII(A)(1)(c), the Approach Query identifies an Approach Event, *i.e.*, a first event, between the first object p and the second object q. Ex. 1003, 62-63. The limitation "first object interacting with the second object" broadly encompasses an embodiment where one object comes within an arbitrary distance, *e.g.*, 3 units, of another object. Ex. 1005, ¶¶252-254. This is consistent with the '923 patent's event embodiment that looks for "two objects com[ing] together." Ex. 1001, 8:62-64. The Approach Query searches the centroid trajectories of the two objects p and q, and the distance between the moving centroids. Ex. 1005, ¶253. The query also searches for trajectories within certain periods, *e.g.*, time-1 for p and time-2 for q. *Id.* The query also applies the object classification attribute, *i.e.*, person, for p and q. *Id.* Because identifying the Approach Event comprises searching for a minimum of two attributes, the first event is not one of the detected attributes. *See* Section VIII(A)(1)(c).

#### g. "wherein the event of the object refers to the object engaged in an activity"

As explained in Section VIII(A)(1)(g), *Kellogg* discloses this element. Ex. 1005, ¶255.

### 6. Independent Claim 9

#### a. "A video device comprising"

The "video device" term does not limit the claim but is merely the intended use of the purported invention, as demonstrated by the substantially corresponding claim limitations between claims 1 and 9. Moreover, the term does not serve as an antecedent basis for limitations in the claim body. Lastly, the prosecution history is devoid of any reliance on the term to distinguish the claims from the prior art.

Nonetheless, *Kellogg* discloses a video device comprising a computer-based visual memory system that includes video cameras, image processing software, a visual memory database, and a user interface as shown in Figure 4-1. Ex. 1003, 68-70, 77-80; Ex. 1005, ¶257.

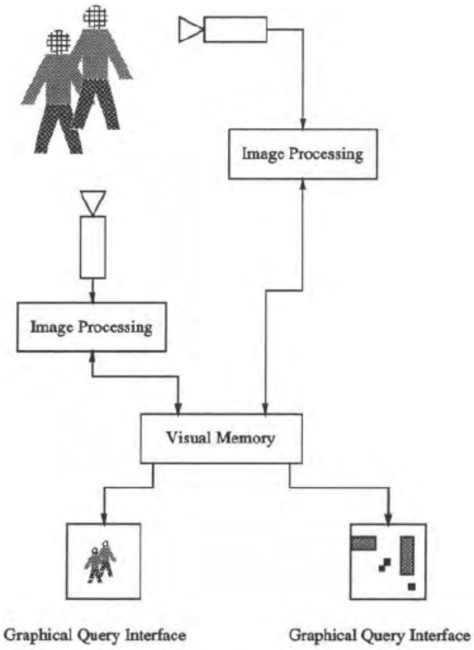



Figure 4-1: Scene monitoring prototype

**b. "means for detecting an object in a video from a single camera"**

*Kellogg* discloses the corresponding structures of this means-plus-function element (*see* Section V(E)(1)). Ex. 1005, ¶¶258-261; *see also* Section VIII(A)(1)(a). *Kellogg's* system is a computer-based system performing the recited functions of these elements. Ex. 1003, 68-70, 77-80. *Kellogg's* system employs the image processing software developed at Texas Instruments for detecting and tracking objects. *Id.*, 77. Indeed, the '923 patent recognizes that any conventional object detection algorithm would work well. Ex. 1001, 9:33-41; Ex. 1005, ¶261.

**c. "means for detecting a plurality of attributes of the object by analyzing the video from said single camera, the plurality of attributes including at least a physical attribute and a temporal attribute, each attribute representing a characteristic of the detected object"**

*Kellogg* discloses the corresponding structures of this means-plus-function element (*see* Section V(E)(2)). Ex. 1005, ¶262; *see also* Section VIII(A)(1)(b). *Kellogg's* system is a computer-based system performing the recited functions of these elements. Ex. 1003, 68-70, 77-80. *Kellogg* reviews several known systems that detect and store attributes of objects. *Id.*, 13-17. Also, *Kellogg's* system relies on the image processing software developed at Texas Instruments for detecting attributes of the object. *Id.*, 77. A POSITA would understand that conventional attribute detection techniques are used in *Kellogg*'s system, much like the '923 patent. Ex. 1001, 10:11-22, 10:27-30, 10:39-41, 10:44-47; Ex. 1005, ¶262.

**d. "a memory storing the plurality of detected attributes"**

As explained in Section VIII(A)(3), *Kellogg* teaches this element. Ex. 1005, ¶263.

**e. "means for selecting a new user rule after the plurality of detected attributes are stored in memory"**

*Kellogg* discloses the corresponding structures of this means-plus-function element (*see* Section V(E)(3)). Ex. 1005, ¶264; *see also* Section VIII(A)(1)(c). The user interface and query mechanisms of *Kellogg*'s system provide the means for selecting a user rule that defines an event. Ex. 1005, ¶264. *Kellogg's* user interface includes a display or monitor (Figs. 4-1, 4-7, 4-8 and related text) and input devices such as a mouse or keyboard for a user to enter queries.

**f. "means for identifying an event of the object that is not one of the detected attributes of the object by applying a selected new user rule to the plurality of attributes stored in memory, for identifying the event independent of when the attributes are stored in memory and for identifying the event without reprocessing the video, wherein the applying the new rule to the plurality of detected attributes comprises applying the new user rule to only the plurality of detected attributes"**

The corresponding structure of this means-plus-function element (*see* Section V(E)(4)) is disclosed by *Kellogg's* query mechanisms as shown in Section VIII(A)(1)(d)-(f). Ex. 1005, ¶265.

**g. "wherein the event of the object refers to the object engaged in an activity"**

As explained in Section VIII(A)(1)(g), *Kellogg* teaches this element. Ex. 1005, ¶266.

**7. Claims 10 and 31**

Claims 10 and 31 depend from claims 9 and 30, respectively. *Kellogg* discloses the added feature of "a video camera operable to obtain the video." Ex. 1003, 68 ("image processing using video ***cameras*** tracks objects and stores information about them in the visual memory" (emphasis added)); Ex. 1005, ¶267.

### 8. Claims 11 and 32

Claims 11 and 32 add that the "means for identifying an event of the [first] object" of claims 9 and 30 comprise "means for identifying a first event [of the object] in real time by analyzing, of the plurality of attributes, only a first selected subset of the plurality of attributes." The corresponding structure of this means-plus-function limitation in claims 11 and 32 (*see* Section V(E)(4)) is the structure identified in Sections VIII(A)(6)(f) and VIII(A)(15)(e), respectively, which performs the real time function. Ex. 1005, ¶¶269-270.

Claim 11 requires "means for identifying a first event of the object" and claim 32 requires "means for identifying an event of the first object interacting with the second object." Both limitations are disclosed as explained in Sections VIII(A)(6)(f) and VIII(A)(15)(f), respectively. Ex. 1005, ¶274.

#### a. "in real time"

The parent claims 9 and 30 require that the "means for identifying" applies the new user rule to the attributes ***stored in memory***. Thus, while claims 11 and 32 recite identifying an object in "real time" it must search stored attributes. Ex. 1005, ¶¶271-272; *see also* Ex. 1001, 9:14-17 ("***archives*** video primitives … and detects event occurrences in ***real time using event discriminators***") (emphasis added).

*Kellogg*'s system allows real-time scene monitoring, and implements real-time operations. Ex. 1003, 10; Fig. 4-8 (annotated); Ex. 1005, ¶273. "[I]t provides the keyword 'now' to signify a ***real-time query***, one that constantly polls the database for new information." *Id.*, 79.

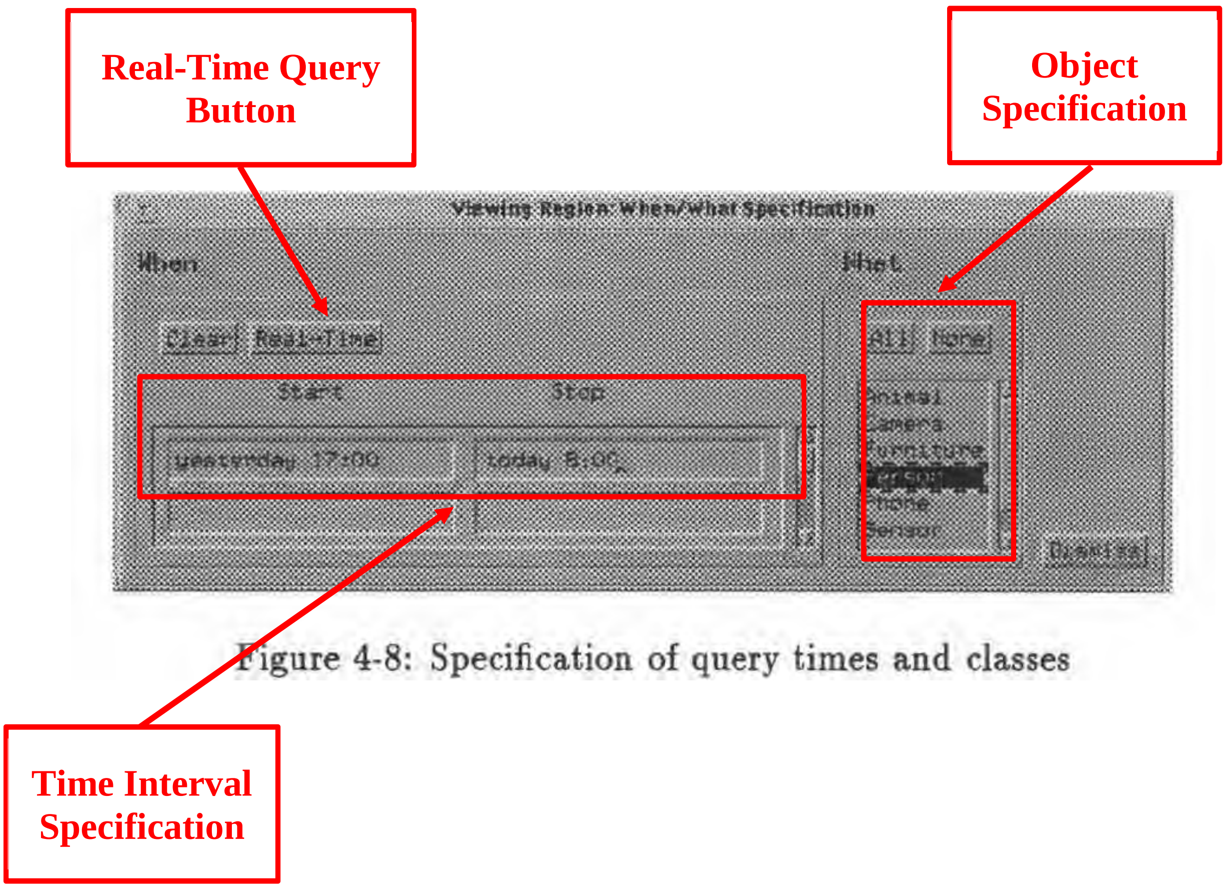



Figure 4-8: Specification of query times and classes

**b. "analyzing, of the plurality of attributes, only a first selected subset of the plurality of attributes"**

*Kellogg*'s Approach Query does not require analyzing all detected attributes as recited as explained in Section VIII(A)(2), therefore it meets the limitation. Ex. 1005, ¶275.

Accordingly, *Kellogg* discloses these claims.

**9. Claims 12 and 33**

Claims 12 and 33 depend from claims 11 and 32, respectively. While claims 11 and 32 recite a means-plus-function limitation requiring the capability of identifying an event relating to an object in ***real-time***, these claims recite a means-plus-function limitation requiring the capability of further identifying an event relating to that same object based on ***stored or archived*** attributes. Ex. 1005, ¶¶276-277. The corresponding structure of this means-plus-function limitation in claims 12 and 33 (*see* Section V(E)(4)) is the structure identified in Sections VIII(A)(6)(f) and VIII(A)(15)(e), respectively, which performs the additional function. Ex. 1005, ¶¶276-277.

*Kellogg* discloses a scene monitoring system that alerts a user of an event in an alarm region. Ex. 1003, 78-79; Ex. 1005, ¶278. The user can set parameters to define queries to detect an event on a user interface. Ex. 1003, Figure 4-8 (annotated). The user can also select a “real-time” mode. *Id.*

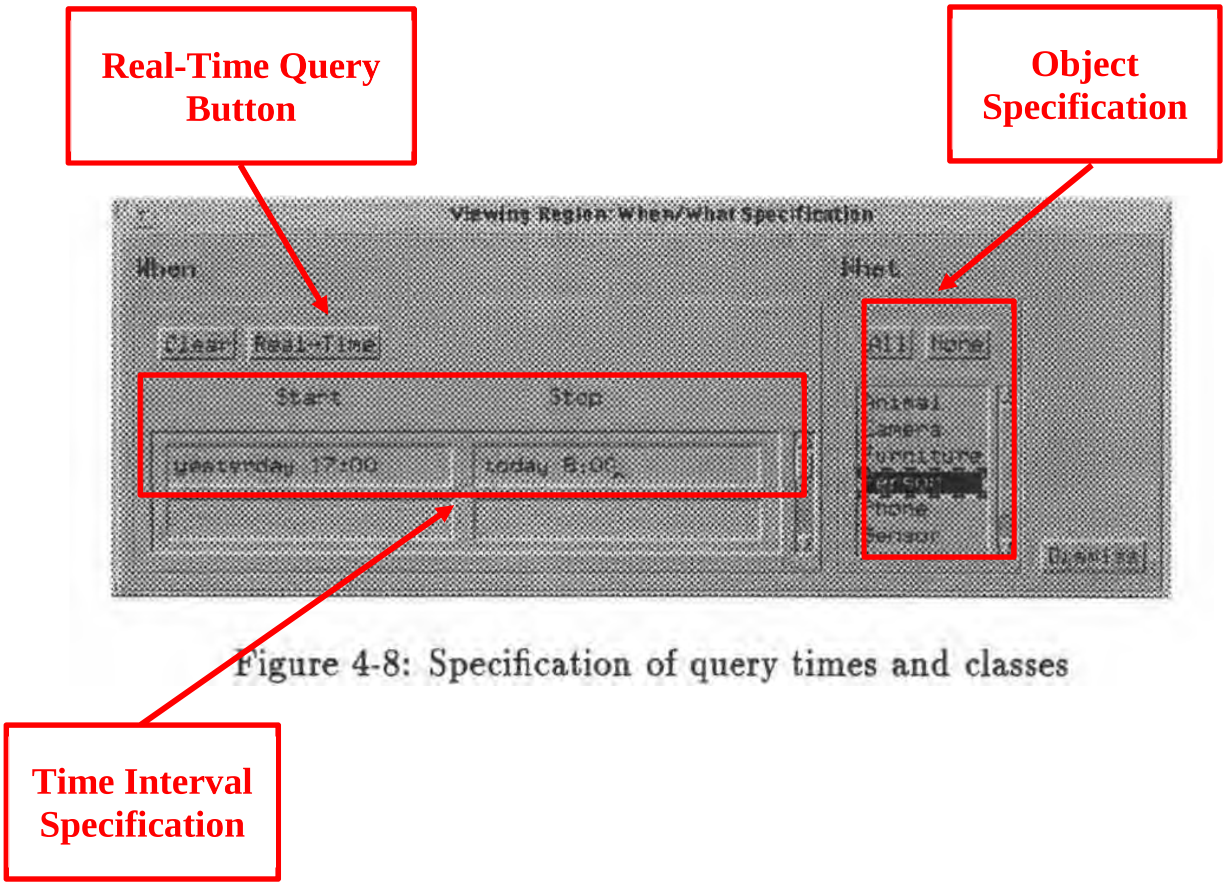



Figure 4-8: Specification of query times and classes

The user can input specific time intervals such as from "yesterday 17:00" to "today 8:00" so that the system searches for ***stored*** or ***archived*** data of the alarm region. *Id.*, Fig. 4-8 (annotated), 79. The system can identify an event relating to the same object identified in real-time mode because the user can select the same object of interest using the Object Specification option. *Id.*, Fig. 4-8 (annotated); Ex. 1005, ¶279. Accordingly, *Kellogg* discloses these claims.

### 10. Claims 14 and 35

Claims 14 and 35 depend from claims 9 and 30, respectively. These claims require the memory being "configured to store at least some of the plurality of attributes for at least two months." These claims are apparatus claims directed to structures, *e.g.*, memories, reciting only a functional requirement, which is met by any conventional non-volatile memory at the priority date. Ex. 1005, ¶281. *Kellogg's* memory would meet this limitation as it is configured to store data indefinitely, and therefore is capable of performing the two-month storage function. *Id.*; Ex. 1003, 54.

*Kellogg* also expressly discloses the additional feature. Ex. 1005, ¶282. *Kellogg's* scene monitoring system allows a user to search past events by entering a specific time interval. Ex. 1003, 79, Fig. 4-8 (annotated). *Kellogg* discloses searching for information from "3/8/93 8:00" until "today 13:00." *Id.*, 79. Based on the publication date of *Kellogg* (May 1993) being two months after "3/8/93," a POSITA would have understood that *Kellogg* teaches the capability of storing attributes for at least two months. Ex. 1005, ¶282.

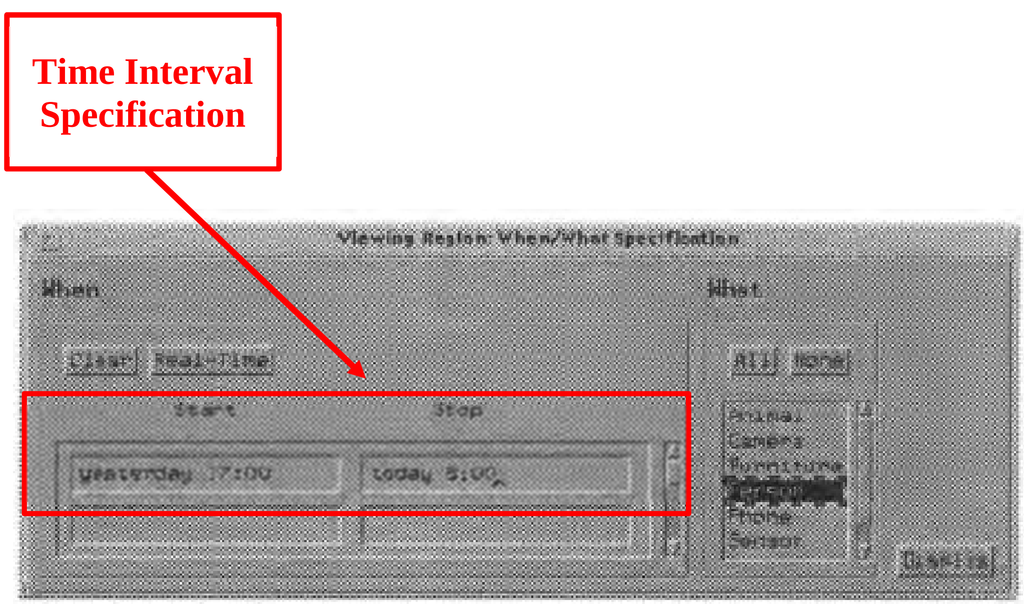



Figure 4-8: Specification of query times and classes

These claims also recite a means-plus-function limitation requiring analyzing some of the attributes that are stored for at least two months. The corresponding structure of this limitation in claims 14 and 35 (*see* Section V(E)(4)) is the structure identified in Sections VIII(A)(6)(f) and VIII(A)(15)(e), respectively, which performs the additional function, thus it is disclosed by *Kellogg*. Ex. 1005, ¶280.

**11. Claims 18 and 40**

Claims 18 and 40 depend from claims 9 and 30, respectively, adding "the video surveillance device is a computer system configured as a video surveillance device." The '923 patent describes, "video surveillance system … is for monitoring a location for … security purposes." Ex. 1001, 4:47-49. *Kellogg* discloses the added feature. Ex. 1005, ¶¶283-284.

*Kellogg's* goal "is to design a visual memory architecture that meets the requirements of ***various computer vision application*** … [and] to implement a visual memory prototype to support a ***real-time scene monitoring prototype***." Ex. 1003, 10 (emphasis added). Further, "the ***scene monitoring system*** alerts the user to events in ***alarm regions*** … [which] can be established all over the map, allowing the user to monitor a number of disjoint regions without having to watch them all." *Id.*, 77-78; 78-79 (a user can set the system to trigger an alarm); 80 (Fig.

4-9 displays one monitored object inside an alarm region and two other monitored objects outside).

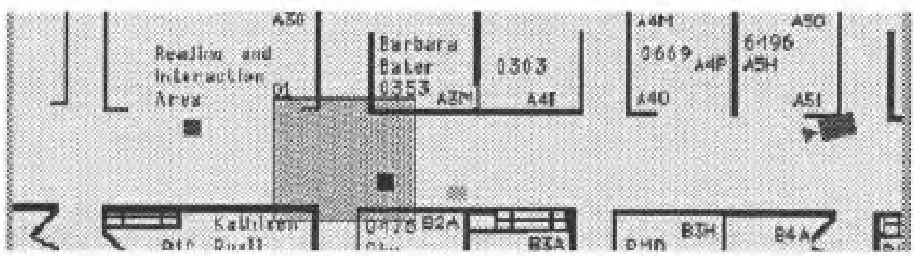



Figure 4-9: Graphical query results

### 12. Claims 19 and 41

Claims 19 and 41 depend from claims 9 and 30, respectively. *Kellogg* discloses “video sensors,” which include video cameras. Ex. 1001, 6:8-12; Ex. 1005, ¶285. *Kellogg*’s system uses more than one video cameras. Ex. 1003, Figure 4-1.

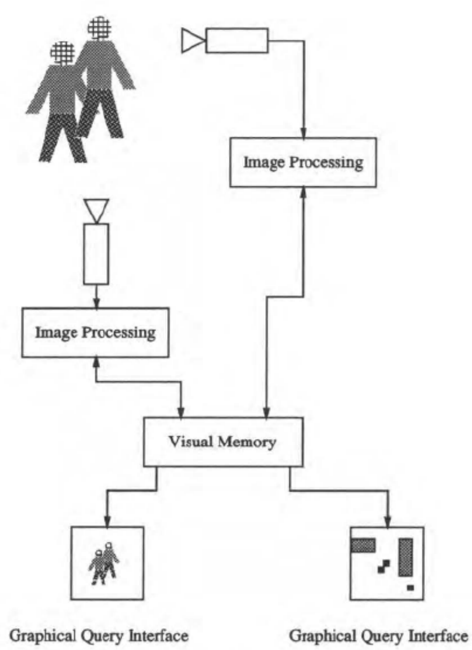



Figure 4-1: Scene monitoring prototype

### 13. Independent Claim 20

The term "video device" does not limit the claim. *See* Section VIII(A)(6)(a). Nevertheless, *Kellogg* discloses this term. *See id.* The rest of claim 20 is substantially identical to claim 1. Ex. 1005, ¶¶286-287. Thus, for the reasons in Sections VIII(A)(1) and VIII(A)(6)(a), *Kellogg* anticipates claim 20.

### 14. Independent Claims 22 and 29

Claims 22 and 29 are essentially claims 1 and 8, respectively, in system form, directed to a "non-transitory computer-readable storage medium." Ex. 1005, ¶289. *Kellogg's* scene monitoring prototype is a software (Ex. 1003, 77), and it is embodied in a computer readable medium, such as memory, hard drive or removable storage media. Ex. 1005, ¶288. Thus, for the reasons in Sections VIII(A)(1) and VIII(A)(5), *Kellogg* anticipates these claims.

### 15. Independent Claim 30

#### a. "A video device comprising"

The term "video device" does not limit the claim. *See* Section VIII(A)(6)(a). Nevertheless, *Kellogg* discloses this term. *See id.*

#### b. "means for detecting first and second objects in a video from a single camera"

As explained in Sections VIII(A)(5)(a) and VIII(A)(6)(b), *Kellogg* discloses this element. Ex. 1005, ¶292.

**c. "means for detecting a plurality of attributes of the object by analyzing the video from said single camera, each attribute representing a characteristic of the respective detected object"**

As explained in Sections VIII(A)(5)(b) and VIII(A)(6)(c), *Kellogg* discloses this element. Ex. 1005, ¶293.

**d. "a memory storing the plurality of detected attributes"**

As explained in Sections VIII(A)(6)(d) and VIII(A)(3), *Kellogg* discloses this element. Ex. 1005, ¶294.

**e. "means for identifying an event of the first object interacting with the second object by applying a selected new user rule to the plurality of attributes stored in memory, and for identifying the event independent of when the attributes are stored in memory, the event not being one of the detected attributes, wherein the applying the selected new user rule to the plurality of attributes stored in memory comprises applying the selected new user rule to only the plurality of attributes stored in memory"**

Compared to claim 9, claim 30 does not recite "and for identifying the event without reprocessing the video," so claim 9 is inclusive. Accordingly, as explained in Sections VIII(A)(5)(c)-(f) and VIII(A)(6)(e)-(f), *Kellogg* discloses this element. Ex. 1005, ¶295.

**f. "wherein the event of the object refers to the object engaged in an activity"**

As explained in Sections VIII(A)(5)(g) and VIII(A)(6)(g), *Kellogg* discloses this element. Ex. 1005, ¶296.

**16. Claim 36**

Claims 36 depends from claim 30 and requires "wherein the means for identifying an event includes means for identifying the event without reprocessing the video." The corresponding structure of this means-plus-function limitation (*see* Section V(E)(4)) is the structure identified in Section VIII(A)(15)(e), performing the additional function. Accordingly, *Kellogg* anticipates claim 36. Ex. 1005, ¶297.

**B. *Kellogg* in view of *Brill* Renders Claims 1-41 Obvious**

A POSITA would have found it obvious to combine *Kellogg*'s system with the teachings of *Brill* to provide enhancements or achieve particular design objectives in *Kellogg*'s system, while yielding predictable results, as explained in Section VI(C), with respect to ***all*** claims.

**1. Claims 1-7**

To the extent Avigilon argues that *Kellogg* does not disclose detecting an object in a video from a "single" camera as required in claim 1, *Brill* teaches such feature as demonstrated in Figure 10. Ex. 1004, 12; Ex. 1005, ¶298.


Figure 10: A series of simple events

If Avigilon argues that *Kellogg*'s disclosure is inadequate or a particular object detection algorithm is required (though not claimed), such as the background-model based technique described in the '923 patent (Ex. 1001, 9:39-48), *Brill* teaches such an object detection technique in detail (Ex. 1004, 6-9). Ex. 1005, ¶299.

If Avigilon argues that the recited "new user rule" requires setting a response, and that *Kellogg* does not disclose it, *Brill* does so because a user can set "actions" such as "beep," "log," or "popup," which are responses to an identified event. Ex. 1004, Figure 11 (annotated); Ex. 1005, ¶300.

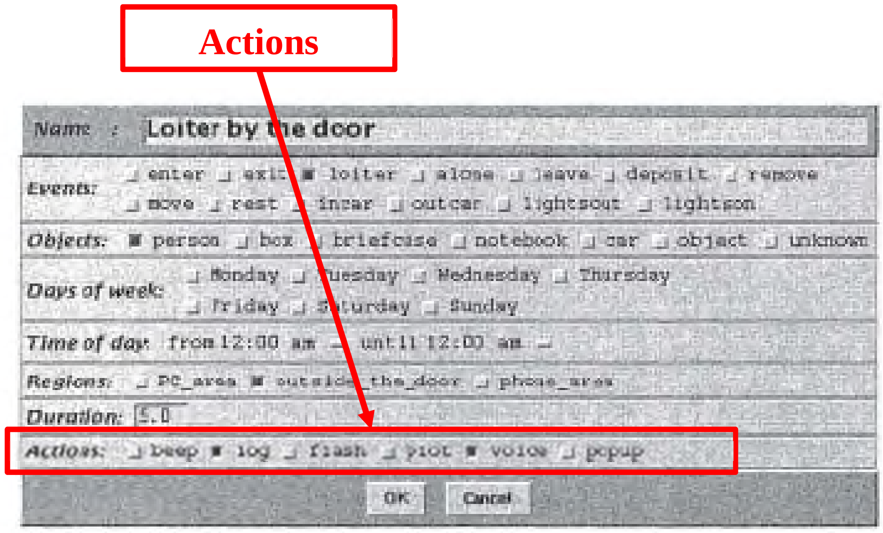



Figure 11: Selecting a type of simple event

### 2. Claim 8

To the extent Avigilon argues that *Kellogg* does not disclose detecting "first" and "second" objects in a video from a "single" camera in claim 8, *Brill* teaches this because its system detects a person (*i.e.*, a first object) and a briefcase (*i.e.*, a second object) in a video obtained from a single camera as demonstrated in Figure 10. Ex. 1004, 12; *see also* 6-9 (detecting people entering and exiting a car); Ex. 1005, ¶302.


Figure 10: A series of simple events

If Avigilon argues that *Kellogg* does not disclose the element of Section VIII(5)(f), *Brill* does. *Brill*'s system identifies a person "picking up" a briefcase, *i.e.*, the first object interacting with the second object, and it analyzes the detected attributes such as the object type (*e.g.*, person, briefcase), activity attributes such as "remove," and the location of each object. Ex. 1004, 12-13; *see also* 6-9; Ex. 1005, ¶304.

If Avigilon argues that the recited "new user rule" requires setting a response, and that *Kellogg* does not, *Brill* does so as explained in Section VIII(B)(1). Ex. 1005, ¶305.

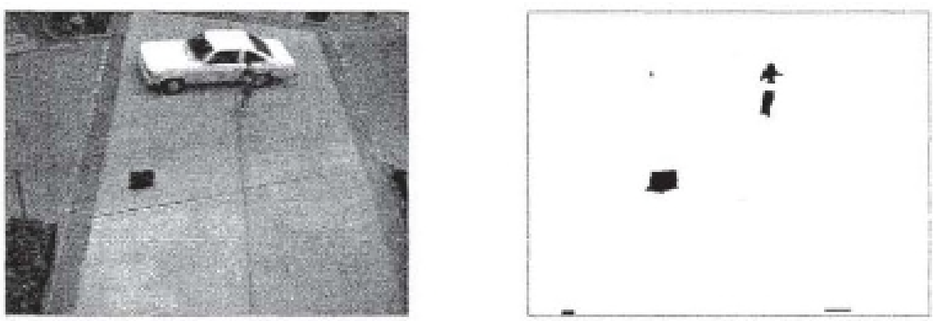
**3. Claims 9-19**

To the extent Avigilon argues that a "video device" is not disclosed in *Kellogg*, *Brill* discloses "smart cameras" wherein the "attributes to be detected are defined in [the] device prior to the selection of a subset of the plurality of attributes." Ex. 1004, 5-6; Ex 1005, ¶307.

If Avigilon argues that *Kellogg* does not disclose the corresponding structure of the limitation "means for detecting an object in a video from a single camera" in claim 9, *Brill* does so. Ex. 1005, ¶¶308-313.

A POSITA would understand the corresponding structure is a computer system or equivalent video processing system utilizing conventional motion and/or change detection algorithms to detect objects. *See* Section V(E)(1); Ex. 1005, ¶309.

*Brill* teaches a "change" detection algorithm that corresponds to "block 52" of the '923 patent. Ex. 1001, 9:39-48 ("[a]ny change detection algorithm" that detects an object "if one or more pixels in a frame are deemed to be in the foreground"). *Brill*'s system employs this algorithm as demonstrated in Figure 4(b) where the person is deemed to be in the foreground. Ex. 1004, 7; Ex. 1005, ¶310.



(a) (b)

Figure 4: (a) Current video image. (b) Foreground difference image

*Brill* also teaches a "motion" detection algorithm that corresponds to "block 51" of the '923 patent. Ex. 1001, 9:33-38 ("Any motion detection algorithm for detecting movement between frames at the pixel level can be used for this block."); Ex. 1005, ¶311. *Brill's* Figure 16 demonstrates that its algorithm detects movement of the two people between frames at the pixel level. Ex. 1004, 15-16;

Ex. 1005, ¶311. "The brightness [of Figs. 16(b) and (d)] indicates the probability that the person's image intersects the given ***pixel***, which is highest in the middle of the region, and falls off towards the edge." Ex. 1004, 15.

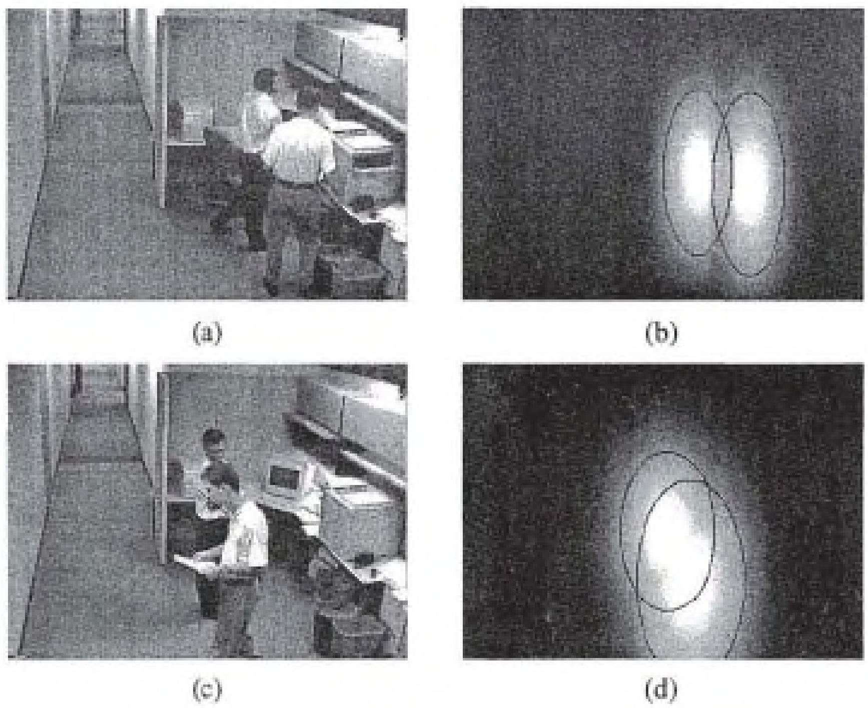


Figure 16: P-template images for partially occluding people

If Avigilon argues that the recited "new user rule" requires setting a response, and that *Kellogg* does not, *Brill* does so as explained in Section VIII(B)(1). Ex. 1005, ¶312.

#### 4. Claims 12 and 33

To the extent Avigilon argues that *Kellogg* does not disclose claims 12 and 33, these claims are obvious over the combination of *Kellogg* and *Brill*. As explained in Section VIII(A)(9), these claims require the capability of identifying a second event relating to the same object that was the subject of a first real-time event, this time based on ***stored*** or ***archived*** attributes.

*Kellogg* discloses this through its discussion of an object identifier, "OID," including searching based on the OID. *See*, Ex. 1003, 50. Thus, once an object is identified in real-time, *Kellogg* could use the OID of the object to perform a second event search on that same object by using the OID. Ex. 1005, ¶317.

*Brill* teaches these claims by explaining that "the system monitors and records the movements of humans in its field of view," and that "[f]or every person that it sees, it creates a log file that summarizes important information about the person, including a snapshot taken when the person was close to the camera and (if possible) facing it." Ex. 1004, 18. "When the person leaves the scene, the log entry is saved to a file. Each log entry records the time when the person entered the scene and a list of coordinate pairs showing their position in each video frame." *Id*.; *see also id*. at Fig. 1 (red).

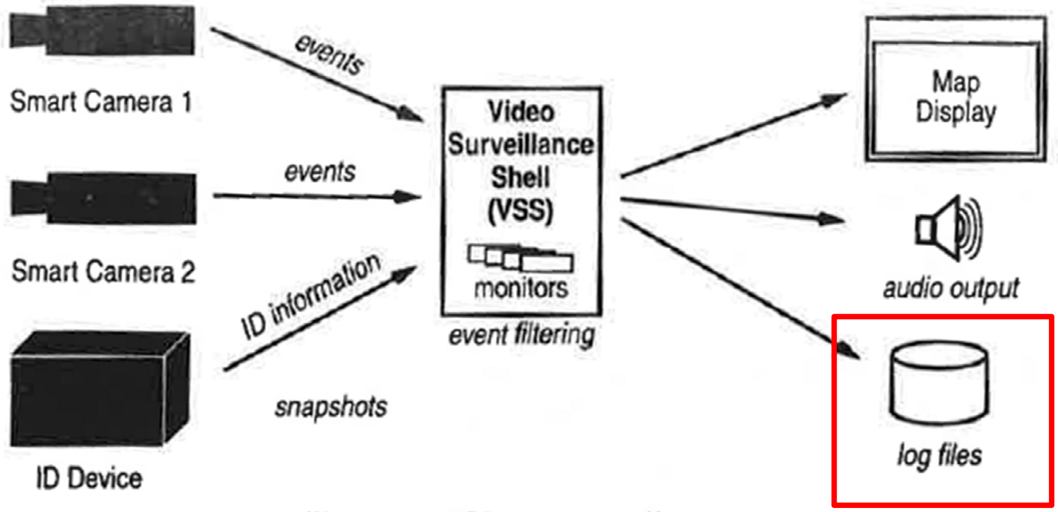



Figure 1: AVS system diagram

*Brill's* queries can run on the entries stored in the "log files," with respect to a particular ID of the object that corresponds to a person of interest that was the subject of a real-time event. *Id*. at 11; Ex. 1005, ¶319.

If Avigilon argues that the recited "second selected subset of the plurality of attributes" must be different from the "first selected subset" of claims 11 and 32 (a proposition that Petitioners does not agree with),[8] a "second event" will necessarily be associated with different attributes compared to the real-time "first event" because the second event will be based on the particular ID of the object. Ex.

[8] The '923 patent specification does not disclose a "second event" or a "second selected subset of the plurality of attributes," much less the "second selected subset of the plurality of attributes" being different from the "first selected subset."

1005, ¶320. Regardless, Petitioners have shown with respect to claims 2 and 4 that *Kellogg* teaches that the query used to determine each particular event can be a subset of attributes that has no relation to any other event. *See* Section VIII(A)(2).

The requirement of claim 33 that the second event relate to the first object interacting with the second object is, as discussed above with respect to claim 32, taught by both *Kellogg* and *Brill*. *See* Section VIII(A)(8).

#### 5. Claims 14 and 35

Claims 14 and 35 depend from claims 9 and 30, respectively. These claims require the memory being "configured to store at least some of the plurality of attributes for at least two months." To the extent Avigilon argues that this feature is not disclosed by *Kellogg*, these claims are obvious over the combination of *Kellogg* and *Brill*. Ex. 1005, ¶¶322-327.

As explained for claims 12 and 33, both *Kellogg* and *Brill* teach identifying events by analyzing stored attributes. *See* Section VIII(B)(4). Neither reference places any limit on how long these attributes can remain in storage prior to their analysis. Ex. 1005, ¶323. This is because the amount of time information is stored is a trivial, non-technical matter of design choice, particularly with respect to an arbitrary, relatively short amount of time, such as two months. *Id.*

Indeed, this limitation requires no technical change to the structures disclosed in *Kellogg* and *Brill. Id.*, ¶325. *Kellogg* teaches a "playback" function where a user can view the monitoring results at a later time. Ex. 1003, 80. *Kellogg* places no time limit on how much later that time could be, and it could be two months, three months, or a year. Ex. 1005, ¶325. Instead, all that is required to meet this limitation is that the user of such a system not delete the stored attributes that are collected for more than two months. *Id.* Of course, conventional computer non-volatile memory is designed to retain the stored information indefinitely. *Id.*, ¶324.

Moreover, the ordinary knowledge of a POSITA would obviously encourage storing attributes for at least two months in a video surveillance system like that disclosed in *Kellogg* and *Brill* if the information is needed for evidence in a prosecution, which would take more than two months. Ex. 1003, 10, 77-78; Ex. 1004, 4-5; Ex. 1005, ¶326; *see* Section VIII(A)(10).

Similarly, *Kellogg* identifies the utility of its system for sports data. Ex. 1003, 53. A POSITA would be motivated to store sports footage attribute data for more than two months, indeed for years. Ex. 1005, ¶327. For example, prior to the priority date of the ’923 patent, Kevin McHale was inducted into the NBA Hall-Of-Fame (Ex. 1039) for his achievements on the court over a decade later. *Id.* Video of these decade old achievements would be of interest to the Hall-Of-Fame committee and fans interested in seeing Mr. McHale at his best. *Id.* This would motivate a POSITA to implement the *Kellogg* system such that the data is stored for well over two months. *Id.* And to take full advantage of this feature, it would be obvious to search some of the attributes stored for at least two months. *Id.*

**6. Claims 18 and 40**

Claims 18 and 40 depend from claims 9 and 30, respectively, and add "the video surveillance device is a computer system configured as a video surveillance device." To the extent Avigilon argues that this feature is not disclosed by *Kellogg*, the claims are obvious over the combination of *Kellogg* and *Brill.* Ex. 1005, ¶¶328-329.

*Brill*'s "Autonomous Video Surveillance (AVS)" system "processes live video streams from surveillance cameras to automatically produce a real-time map-based display of the locations of people, objects and events in a monitored region." Ex. 1004, 4; *see also id.*, 5 ("AVS system incorporates multiple cameras to enable surveillance"). The Video Surveillance Shell (VSS) in Figure 1 is a computer system that integrates the information from the surveillance cameras and displays it on a map. *Id.*, 4; Ex. 1005, ¶329. A POSITA would have readily configured *Kellogg's* system to include teachings of *Brill* because *Kellogg* also contemplates its system being used in scene monitoring. Ex. 1003, 77-80; Ex. 1005, ¶329; *see* Section VIII(A)(10).

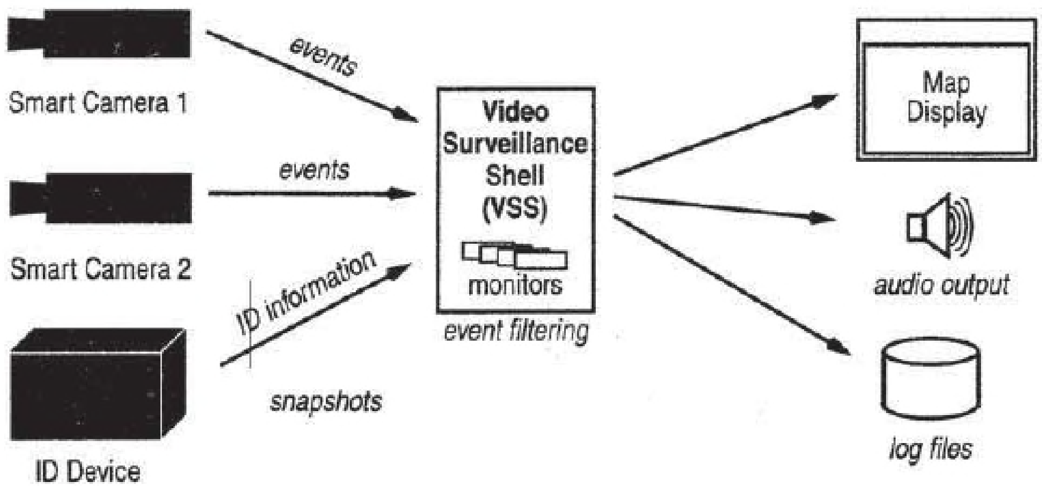



Figure 1: AVS system diagram

### 7. Claims 20-21

To the extent Avigilon argues that a "video device" is not disclosed in *Kellogg*, *Brill* explicitly discloses "smart cameras" wherein the "attributes to be detected are defined in [the] device prior to the selection of a subset of the plurality of attributes." Ex. 1004, 5-6; Ex. 1005, ¶330.

If Avigilon argues that *Kellogg* does not disclose "selecting the new user rule to provide an analysis of a ***combination*** of the ***plural*** physical attributes and the ***plural*** temporal attributes to detect the event" (claim 21), *Brill* explicitly discloses this because a user can select a "region" and "object type" (*i.e.*, plural physical attributes), and a "day of week" and "time of day" (*i.e.*, plural temporal attributes). Ex. 1004, Figure 11; Ex. 1005, ¶331.

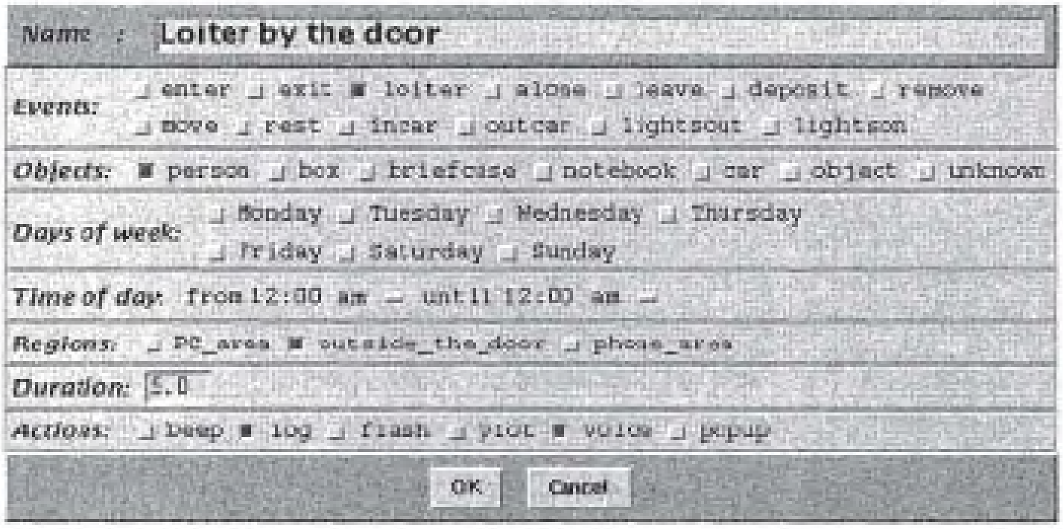



Figure 11: Selecting a type of simple event

If Avigilon argues that the recited "new user rule" requires setting a response, and that *Kellogg* does not, *Brill* does so as explained in Section VIII(B)(1). Ex. 1005, ¶332.

**8. Claims 22-29**

Claims 22 and 29 are essentially claims 1 and 8, respectively, in system form, directed to a "non-transitory computer-readable storage medium." To the extent Avigilon argues *Kellogg* lacks a disclosure of a "non-transitory computer-readable storage medium," *Brill* teaches it. Ex. 1005, ¶¶334-335. *Brill's* AVS system is a "software" (Ex. 1004, 6) that is embodied in a computer readable medium, such as memory, hard drive or removable storage media. Ex. 1005, ¶335.

**9. Claims 30-41**

To the extent Avigilon argues that a "video device" is not disclosed in *Kellogg*, *Brill* discloses this. *See* Section VIII(B)(3); Ex. 1005, ¶336.

If Avigilon argues that *Kellogg* does not disclose the corresponding structure of the limitation "means for detecting an object in a video from a single camera," *Brill* does so. *See* Section VIII(B)(3); Ex. 1005, ¶¶337-340.

If Avigilon argues that the recited "new user rule" requires setting a response, and that *Kellogg* does not, *Brill* does so. *See* Section VIII(B)(1). Ex. 1005, ¶341.

If Avigilon argues that *Kellogg* does not disclose the element of Section VIII(15)(e), *Brill* does do. *See* Section VIII(B)(2); Ex. 1005, ¶¶342-343.

# IX. MANDATORY NOTICES UNDER 37 C.F.R. § 42.8

## A. Real Party-in-Interest

Petitioner certifies that the real parties-in-interest are Canon Inc., Canon U.S.A., Inc., and Axis Communications AB, and that no other party exercised control or could exercise control over Petitioner's participation in this proceeding, the filing of this Petition, or the conduct of any ensuing trial.

## B. Related Matters

Concurrently with this Petition, Petitioner is filing a separate Petition for IPR for the '923 patent, and two separate Petitions for IPRs for the '912 patent. IPR2018-00138 and IPR2018-00140 for the '661 patent are related matters and were filed on October 31, 2017, instituted on June 1, 2018, and are currently pending.

### C. Lead and Back-Up Counsel and Service Information

Petitioners Canon Inc. and Canon U.S.A., Inc. provide the following designation of counsel:

| **Canon Inc. and Canon U.S.A., Inc.<br>Lead Counsel** | **Canon Inc. and Canon U.S.A., Inc.<br>Back-Up Counsel** |
|---|---|
| Joseph A. Calvaruso<br>Reg. No. 28,287<br>Orrick, Herrington & Sutcliffe LLP<br>51 West 52nd Street<br>New York, NY 10019-6142<br>Telephone: (212) 506-5140<br>Fax: (212) 506-5151<br>JVCPTABDocket@orrick.com | Richard F. Martinelli<br>Reg. No. 52,003<br>Orrick, Herrington & Sutcliffe LLP<br>51 West 52nd Street<br>New York, NY 10019-6142<br>Telephone: (212) 506-3702<br>Fax: (212) 506-5151<br>RFMPTABDocket@orrick.com |

Petitioner Axis Communications AB provides the following designation of counsel:

| **Axis Communications AB<br>Lead Counsel** | **Axis Communications AB<br>Back-Up Counsel** |
|---|---|
| C. Gregory Gramenopoulos<br>Reg. No. 36,532<br>Finnegan, Henderson, Farabow,<br>Garrett & Dunner LLP<br>901 New York Avenue, NW<br>Washington, DC 20001-4413<br>Tel. 202.408.4263; Fax 202.408.4400<br>gramenoc@finnegan.com | Jessica L.A. Marks<br>Reg. No. 67,451<br>Finnegan, Henderson, Farabow,<br>Garrett & Dunner LLP<br>Two Freedom Square<br>11955 Freedom Drive<br>Reston, VA 20190-5675<br>Tel. 571.203.2791; Fax 202.408.4400<br>jessica.marks@finnegan.com |

In accordance with the above designations, Powers of Attorneys are submitted with this Petition. Please address all correspondence to lead and back-up counsel. Petitioner consents to service by email at the addresses listed in the above tables.

## X. CONCLUSION

For the reasons presented in this Petition, challenged claims 1-41 are unpatentable. Petitioner respectfully requests that the Board grant IPR of claims 1-41 and institute trial.

Dated: November 12, 2018

By: /C. Gregory Gramenopoulos/

C. Gregory Gramenopoulos
Reg. No. 36,532

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER LLP

901 New York Avenue, NW
Washington, DC 20001-4413
Tel. 202.408.4263; Fax 202.408.4400

gramenoc@finnegan.com

*Attorney for Petitioner*
*Axis Communications AB*

By: /Joseph A. Calvaruso/

Joseph A. Calvaruso
Reg. No. 28,287

ORRICK, HERRINGTON & SUTCLIFFE LLP

51 West 52nd Street
New York, NY 10019-6142
Tel. 212.506.5140; Fax 212.506.5151

JVCPTABDocket@orrick.com

*Attorney for Petitioners*
*Canon Inc. and Canon U.S.A., Inc.*

**CERTIFICATE OF SERVICE**

The undersigned hereby certifies that a copy of the foregoing Petition for IPR of U.S. Patent No. 7,932,923, including all Exhibits and the Power of Attorney, was served on November 12, 2018 via FedEx, overnight delivery, directed to the attorney of record for the patent at the following address:

Avigilon Corporation
Box 378, #101-1001 West Broadway
Vancouver BC V6H 4E4

Dated: November 12, 2018

By: /Joseph A. Calvaruso/

Joseph A. Calvaruso
Reg. No. 28,287

ORRICK, HERRINGTON & SUTCLIFFE LLP

51 West 52nd Street
New York, NY 10019-6252
Telephone: (212) 506-5140
Fax: (212) 506-5151
JVCPTABDocket@orrick.com

*Attorney for Petitioners*
*Canon Inc. and Canon U.S.A., Inc.*

**CERTIFICATE OF WORD COUNT COMPLIANCE**

Pursuant to 37 C.F.R. § 42.24(d), the undersigned certifies that the foregoing Petition, including footnotes, contains as measured by the word-processing system used to prepare this paper, 13,976 words. This word count does not include the items excluded by 37 C.F.R. § 42.24(a).