PATROLL Winning Submission

U.S. Patent 7,233,684

U.S. Patent 7,233,684 ("*Monument Peak*" or the "patent-at-issue") was filed on November 25, 2002. Claim 1 of the patent-at-issue is generally directed to an imaging method where a scene is being captured and affective information is collected. The affective information is based on the physiological characteristics of the person upon seeing the image and the degree of importance is associated with the captured image.

The primary reference, U.S. Patent 8,561,095 ("*Sisvel*"), was filed on November 13, 2001, and claims priority on the same date. The patent is directed to a method of analyzing movements and reactions of a person watching a program from a television. Sensors monitor the viewers for an emotional response that can be associated with the program being displayed. Responses are being analyzed and the system can monitor and recommend programs based on the viewer's preference.

A sample claim chart comparing claim 1 of *Monument Peak* to *Sisvel* is provided below.

US7233684 ("*Monument Peak*")	A. US8561095 ("*Sisvel*")
1.pre. An **imaging method** comprising the steps of:	**A. US8561095** "5. The system of claim 1, wherein the at least one sensor further comprises **a video image capturing device**." *Sisvel* at claim 5 "14. The method of claim 9, wherein the monitoring is performed by **an image-capturing device for capturing images of the viewer viewing the programming**." *Sisvel* at claim 14 "**Video camera 240 may be used to provide sensory information in a variety of ways**. Unlike most of the other sensors previously recited (with the exception of the microphone), its signal output can merely be saved onto a video recording device such as video recorder 140 shown in FIG. 1. Normally, however, a separate video recorder (not shown) will be used if the viewer-reaction video images are to be recorded as such." *Sisvel* at col. 7:22-28
1.a. **capturing an image of a scene**;	**A. US8561095** "14. The method of claim 9, wherein the monitoring is performed by **an image-capturing device for capturing images of the viewer viewing the programming**." *Sisvel* at claim 14 "**Video camera 240 is positioned to capture an image of a viewer seated in chair 215, and in particular, the head and face region of the viewer**. It may be stationary or movable." *Sisvel* at 7:16-18 "FIG. 3 is a block diagram illustrating a video processor 300 that may be used to process input from a video camera 240 according to a preferred embodiment of the present invention. Video processor 300 receives video input from camera 240 at input port 305. Preferably, video processor 300 also includes video recording medium 310 for **recording the image captured by camera** 240." *Sisvel* at 7:49-55

1.b. **collecting affective information at capture**; and	**A. US8561095** "1. A system comprising: . . . at least one sensor for **sensing a physical reaction** by a **viewer viewing the displayed programming and generating a signal representative of the physical reaction**; . . . ." *Sisvel* at claim 1 "FIG. 2 illustrates an exemplary system for **collecting viewer reactions to programs and program segments** that is implemented in a system analogous to the exemplary television receiver system 100 of FIG. 1, according to a similar embodiment of the present invention." *Sisvel* at 6:7-11
1.c. **associating the affective information with the scene image**, wherein the step of **collecting affective information comprises monitoring the physiology of a user** and, wherein the step of collecting affective information comprises the steps of **interpreting the collected physiological information to determine the relative degree of importance of the scene image**.	**A. US8561095** "1. A system comprising: . . . **a processor for receiving and analyzing the signal representative of the physical reaction to determine if it can be associated with at least one recognizable viewer emotional response**, whereupon satisfying said determination **the processor associates the at least one recognizable viewer response with descriptive information that was being displayed** during a time-stamped program segment of the displayed programming when **the physical reaction was sensed by the at least one sensor**, to provide a viewer preference relating to the descriptive information; . . . **a monitoring system that monitors subsequent programs that become available to be displayed**; **a recommender that is configured to recommend subsequent programs based on the viewer preference**; and . . . ." *Sisvel* at claim 1 "9. A method of **assessing viewer response to television programming** that includes one or more distinct timed segments, said method comprising: . . . **monitoring at least one viewer physical condition**; **associating the at least one viewer physical condition with a viewer response**; . . . ." *Sisvel* at claim 9 "FIG. 4 is a system diagram illustrating the interaction of sensory inputs according to one embodiment of the present invention, as well as possible applications for using the

(cont.) 1.c. **associating the affective information with the scene image**, wherein the step of **collecting affective information comprises monitoring the physiology of a user** and, wherein the step of collecting affective information comprises the steps of **interpreting the collected physiological information to determine the relative degree of importance of the scene image**.	gathered and processed sensory data. Again, this embodiment is illustrative and not limiting. As described above, **facial expressions and head movements are analyzed against reference images to determine recognizable emotions** (block 405). **Indicators such as smiles and the intensity of a gaze** (that is, duration without moving and the direction of gaze) **are interpreted** according to a set of predetermined criteria. **Gross actions are also interpreted, such as movement around the room** (block 410) as detected by video camera 240 or motion sensor 230. In addition, **biometric response data captured by sensors** on, for example, chair 215 or sensor shirt 220 (block 415) **may be characterized as associated with certain human emotional responses** in a manner similar to that described above in association with captured video images, albeit with somewhat less certainty in most cases." *Sisvel* at col. 8:23-41 "In a preferred embodiment, the **extracted facial expressions from the visual domain are all labeled by symbols and given values for the feature vector**. For example, **if there are five states: sad, laughing, indifferent, bored, afraid, then these are symbolically encoded** (preferably with numbers). Also, **the volume of the emotion is recorded (e.g. from mild to extreme happiness) on a scale from 0 to 1, and the degree of certainty of the algorithm is also recorded**. The **viewer's pose and actions are tracked and these are also recorded: sitting, getting up, distracted reading a newspaper, shouting, etc.** (these are also specific values for our feature vector)." *Sisvel* at col. 9:15-25 "The augmenting material is presumably available from the programming provider, but used according to the individual tastes of the viewer. For example, **a viewer who has previously demonstrated a preference for action scenes may have a car chase inserted in place of a courtroom scene, while another viewer who prefers comedy may have a humorous sequence instead**." *Sisvel* at col. 11:47-53