Healthcare Support through Information Technology Enhancements (hSITE)

Task 2.2.1 **Sensor Information Acquisition and Feature Extraction for the Real-Time Interpretation of Patient and Workflow Information from Video Feeds**,

Task 2.2.3 **Context Aware Multimodal Information Fusion.**

Emil M. Petriu

School of Information Technology and Engineering
University of Ottawa

Annual Research Review
June 4th, 2010

Theme 2: Context Aware Sensors Systems, Software and Applications

Project 2.1: Protocols, Software Engineering and Software Architectures

Task 2.1.1: Software Engineering for Clinical Context Aware Services

Task 2.1.2: Middleware and Supporting Software

Project 2.2: Information Gathering from the Complex Multi-Sensor Environment

Task 2.2.1: Sensor Information Acquisition and Feature Extraction

Task 2.2.2: Advanced Compression and Fidelity/Rate Modeling for Resource Allocation

Task 2.2.3: Context Aware Multimodal Information Fusion

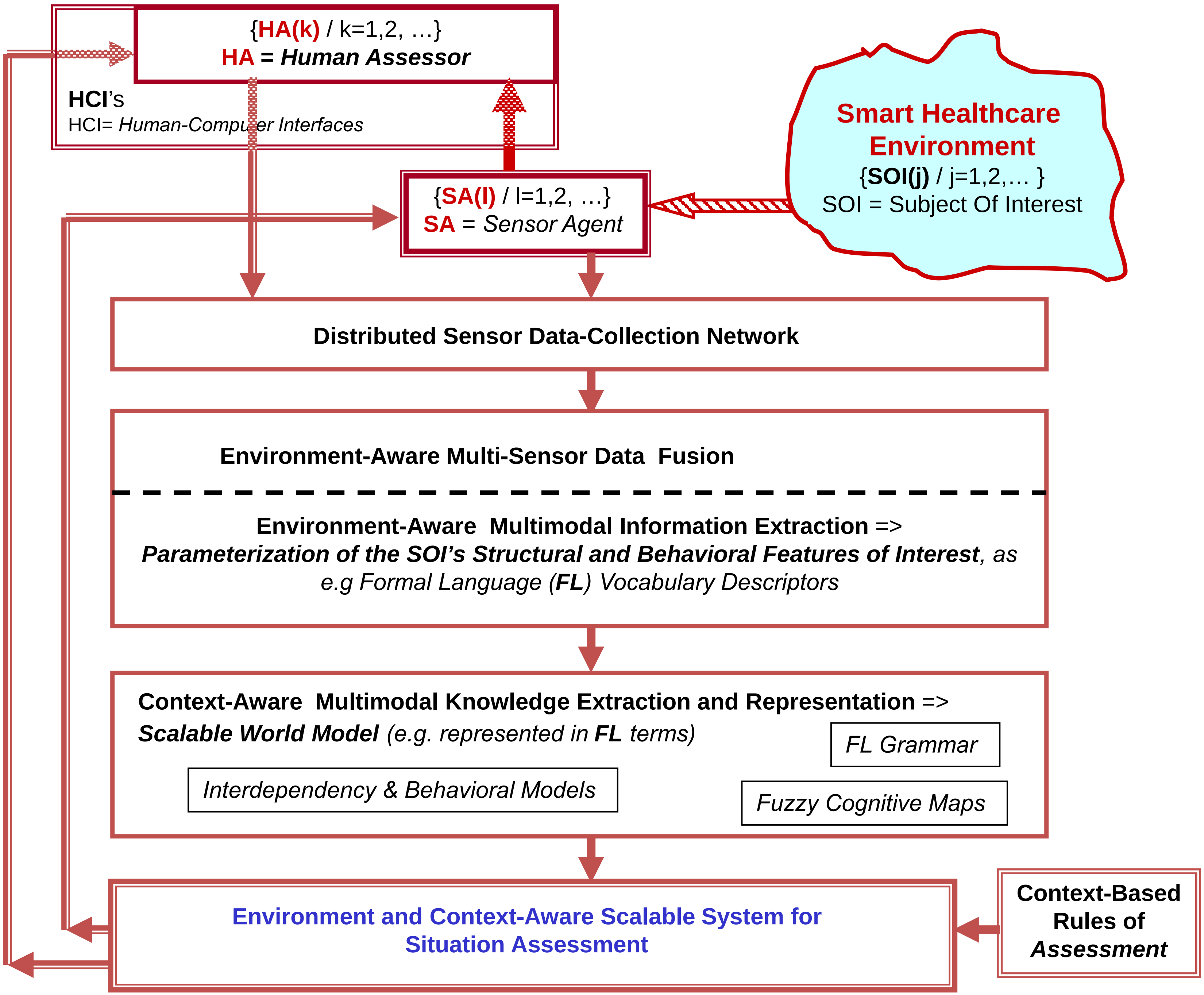
Our team is working on the development of Clinical Care Grade Solutions for Information Gathering from the Complex Multi-Sensor Environments (*Project 2.2*) for **Critical Care** (*Context* #1) and **Home Care** (*Context* #2) applications.

Research Group

Name of student/PDF	Program MSc/PhD/ PDF/ u-grad	Task #	hSITE Start date	Graduation date (actual or anticipated)	Funded by:
Ana-Maria Cretu,	PDF	2.2.1. & 2.2.3.	Jan. 2009	-	hSITE & other
Duo Liu	M.Eng.	2.2.1.	Sept. 2009	April. 2010	hSITE & other
Qing Chen	PDF	2.2.1.	Oct. 2008	-	other
Albino Cordeiro	PDF	2.2.3.	Jan 2009	-	other
Marius D. Cordea	Research Associate	2.2.1.	Oct 2009	Dec.2009	other

Smart-healthcare environments incorporate a multitude of *time-* and *location-dependent sensor-data*, from which is possible to extract relevant information about *patient condition* (identity, location, physiological parameters), *clinical staff status* (identity, location, readiness), *specific clinical activities, medication, supplies, and equipment status* (identity, location, specs), *operating room readiness, state of the ambient environment*, etc.

Context understanding in these environments require dynamic sensor configurations and measurement capabilities similar to human perception, which pose a considerable challenge to the traditional sensor fusion methods. *Location, together with time, represents one of the basic contextual information to any context-aware system*.


{HA(k) / k=1,2, …}
HA = Human Assessor
HCI's
HCI= Human-Computer Interfaces
Smart Healthcare Environment
{SOI(j) / j=1,2,… }
SOI = Subject Of Interest
{SA(l) / l=1,2, …}
SA = Sensor Agent
Distributed Sensor Data-Collection Network
Environment-Aware Multi-Sensor Data Fusion
Environment-Aware Multimodal Information Extraction =>
Parameterization of the SOI's Structural and Behavioral Features of Interest, as e.g Formal Language (FL) Vocabulary Descriptors
Context-Aware Multimodal Knowledge Extraction and Representation =>
Scalable World Model (e.g. represented in FL terms)
FL Grammar
Interdependency & Behavioral Models
Fuzzy Cognitive Maps
Environment and Context-Aware Scalable System for Situation Assessment
Context-Based Rules of Assessment

We are concentrating on two tasks:

Task 2.2.1 **Sensor Information Acquisition and Feature Extraction for the Real-Time Interpretation of Patient and Workflow Information from Video Feeds**,

We are studying and evaluating

(i) different video-data acquisition, image processing and computer vision techniques to be used for the body posture tracking and recognition from video-data streams;

(ii) the **IBM Sensor Event Platform for Healthcare and** the **IBM Smart Sensor Solutions** as ***industrial strength sensor data acquisition platforms***.

Task 2.2.3 **Context Aware Multimodal Information Fusion.**

We are studying and evaluating: a multi-sensor fusion system architecture based on the mission-critical JDL Data Fusion Model developed by the US DoD -Data Fusion Group, able to manage in a consistent multiple sensor data streams;

The **multi-sensor fusion architecture** will be based on the mission-critical *JDL Data Fusion Model* developed by the Joint Directors of Laboratories Data Fusion Group .

This architecture has five functional levels.

* *level 0* "Signal/Feature Assessment" and *level 1* "Entity Assessment" essentially asses the measurement data;

* *level* 2 "Situation Assessment";

* *level 3* "Impact Assessment" essentially asses the information recovered from data;

* *level 4* "Performance Assessment" provides sensor management functions for process refinement.

* a supplementary knowledge-management *level 5* "User Refinement" is used delineate the human from the computer in the process refinement and allow for the adaptive decision of who can query and respectively access the information and the collected data in order to support cognitive decision support and actions.

We **met the objectives for the following milestones**:

- [M2.7.a] Preliminary investigation of body posture recovery algorithms for video interpretation (Task 2.2.1)
- [M2.12] Initial design of a JDL-type multi sensor data fusion system framework (Task 2.2.3)

We are **currently working on the following milestones**:

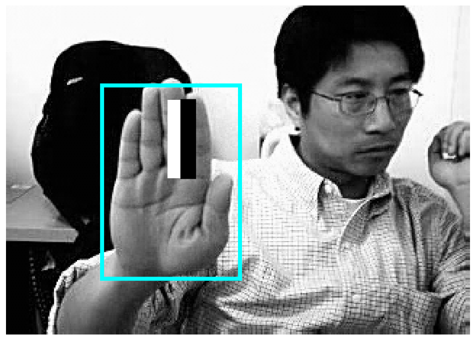
- [M2.7.b] Development and implementation of real-time body posture recovery algorithms for videointerpretation (Task 2.2.1)
- [M2.13a] Study of fuzzy and stochastic information fusion techniques as aids to clinicians (Task2.2.3)

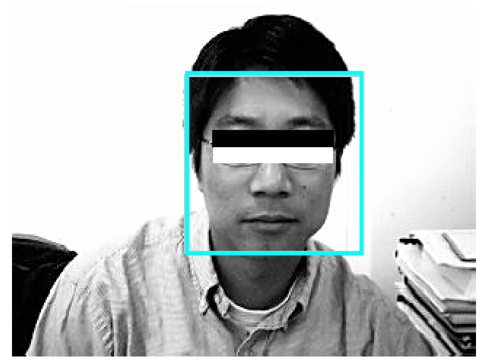
Research Work Accomplished

- Working on the **development of algorithms for the video body posture recovery, under [M2.7.a] and [M2.7.b**], we investigated the real-time performance and the portability/extension-potential of the image processing algorithms that our team developed for *hand-gesture and facial-expression tracking and recognition*, [Chen2009a], [Chen2009b].

Hand Gesture Recognition

- To detect the hand, the image is scanned by a sub-window containing Haar-like feature.



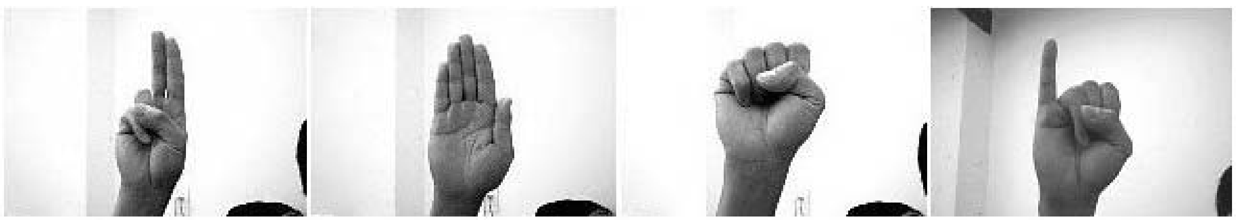


- Based on each Haar-like feature f_j , a weak classifier $h_j(x)$ is defined as:

$$h_j(x) = \begin{cases} 1 & if \quad p_j f_j(x) < p_j \theta_j \\ 0 & otherwise \end{cases}$$

where x is a sub-window, and θ is a threshold. p_j indicating the direction of the inequality sign.

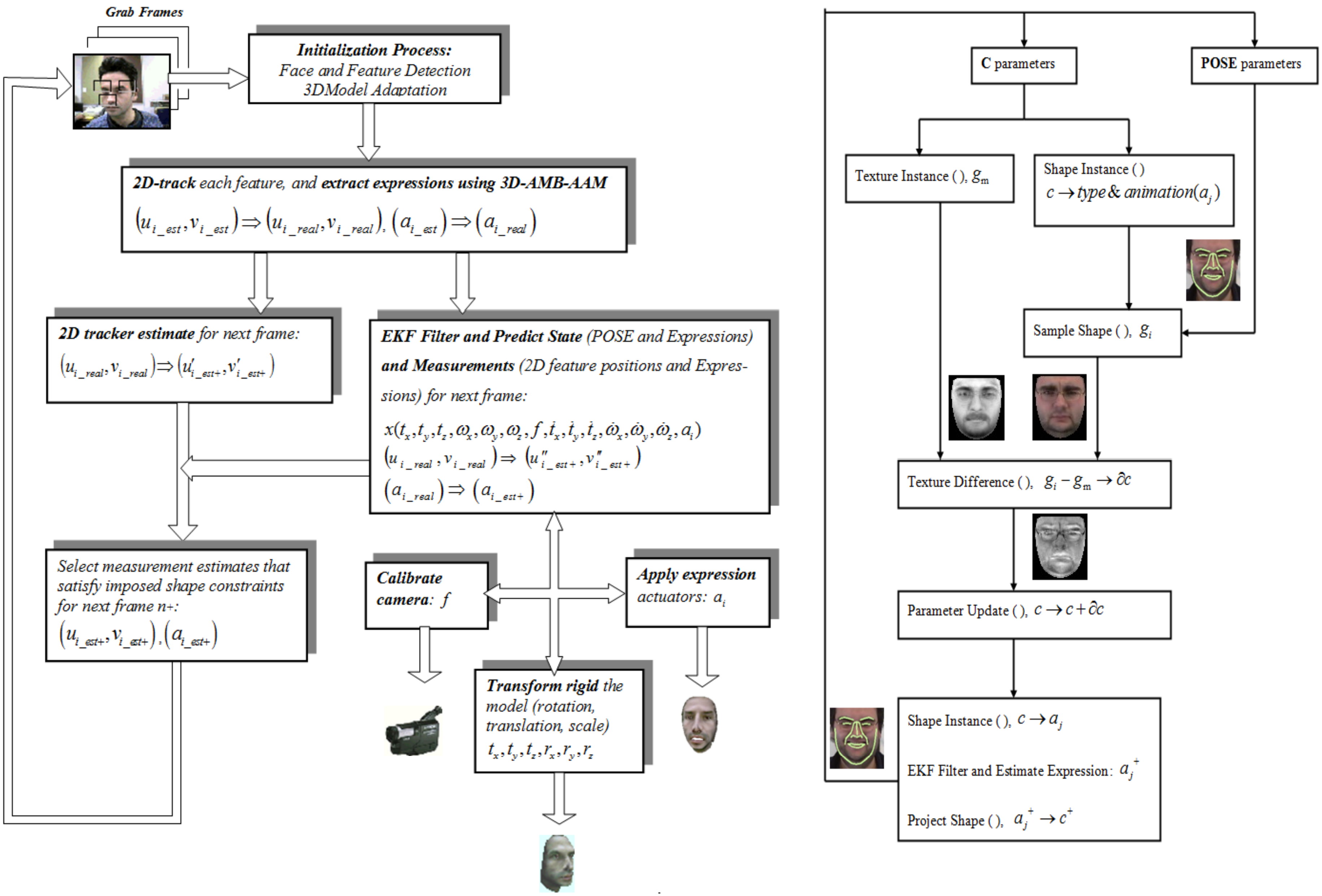
- Four hand postures have been tested with Viola & Jones algorithm:



(a) two-finger (b) palm (c) fist (d) little finger

- Input device: A low cost Logitech QuickCam web-camera with a resolution of 320 $\times$ 240 up at 15 frames-per-second.

Tracker Implementation



Real-time EKF tracking

Model-Based Facial Expression Recognition

- Person Dependent

AU	Signification	No.	Correct	False	Missed	Confused	Recognition Rate
0	Neutral	20	19	1	0	0	95%
1	Inner Brow Raiser	24	21	0	3	0	87.5%
2	Outer Brow Raiser	52	41	1	9	1	78.8%
4	Brow Lowerer	42	41	0	0	1	97.6%
12	Lip Corner Puller	51	48	3	0	0	94.1%
15	Lip Corner Depressor	18	17	0	1	0	94.4%
26	Jaw Drop	74	55	0	19	0	74.3%
Total		281	242	5	32	2	86.1%
False Alarm: 1.7%, Missed: 11.3%							

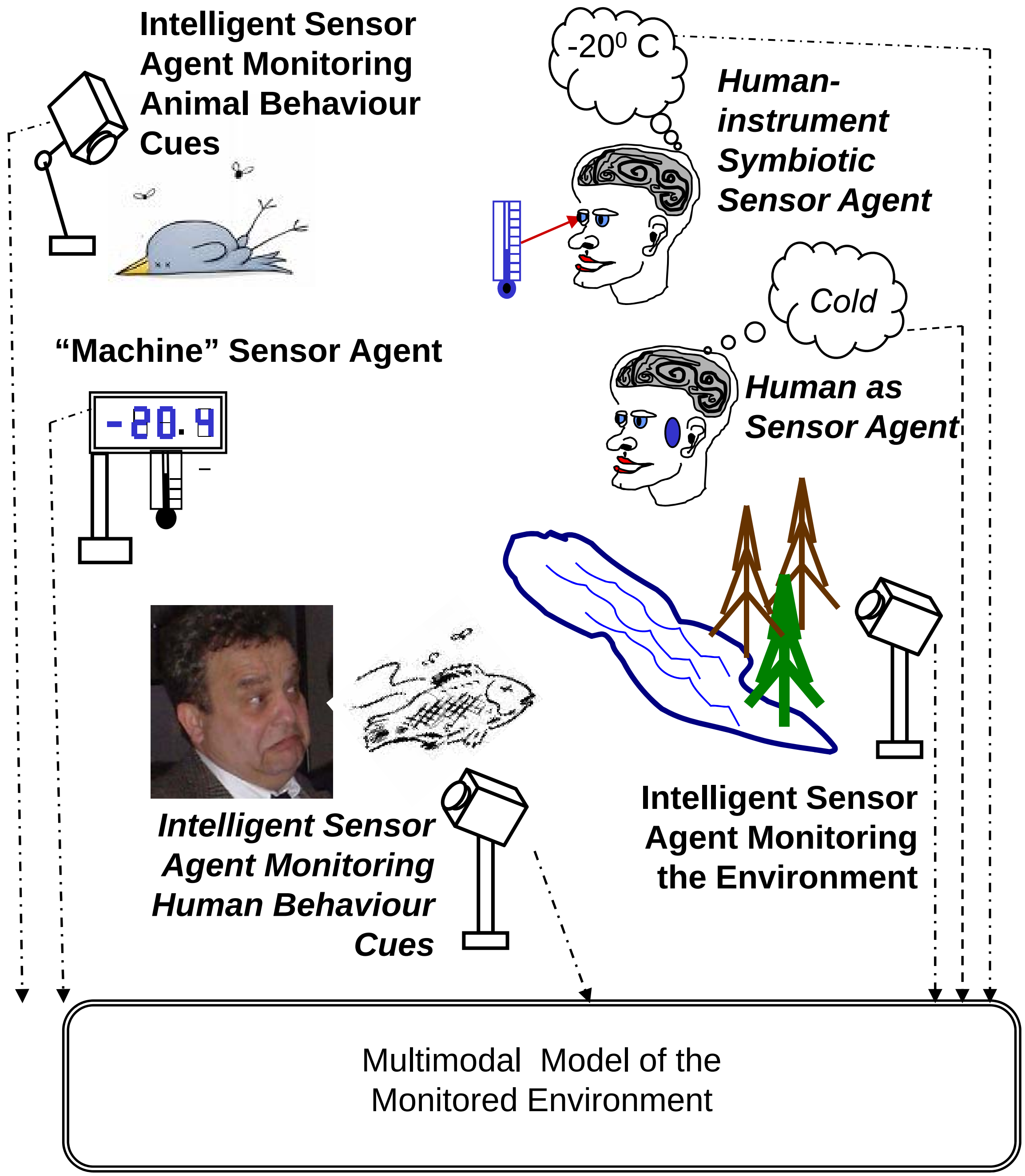
- Person Independent

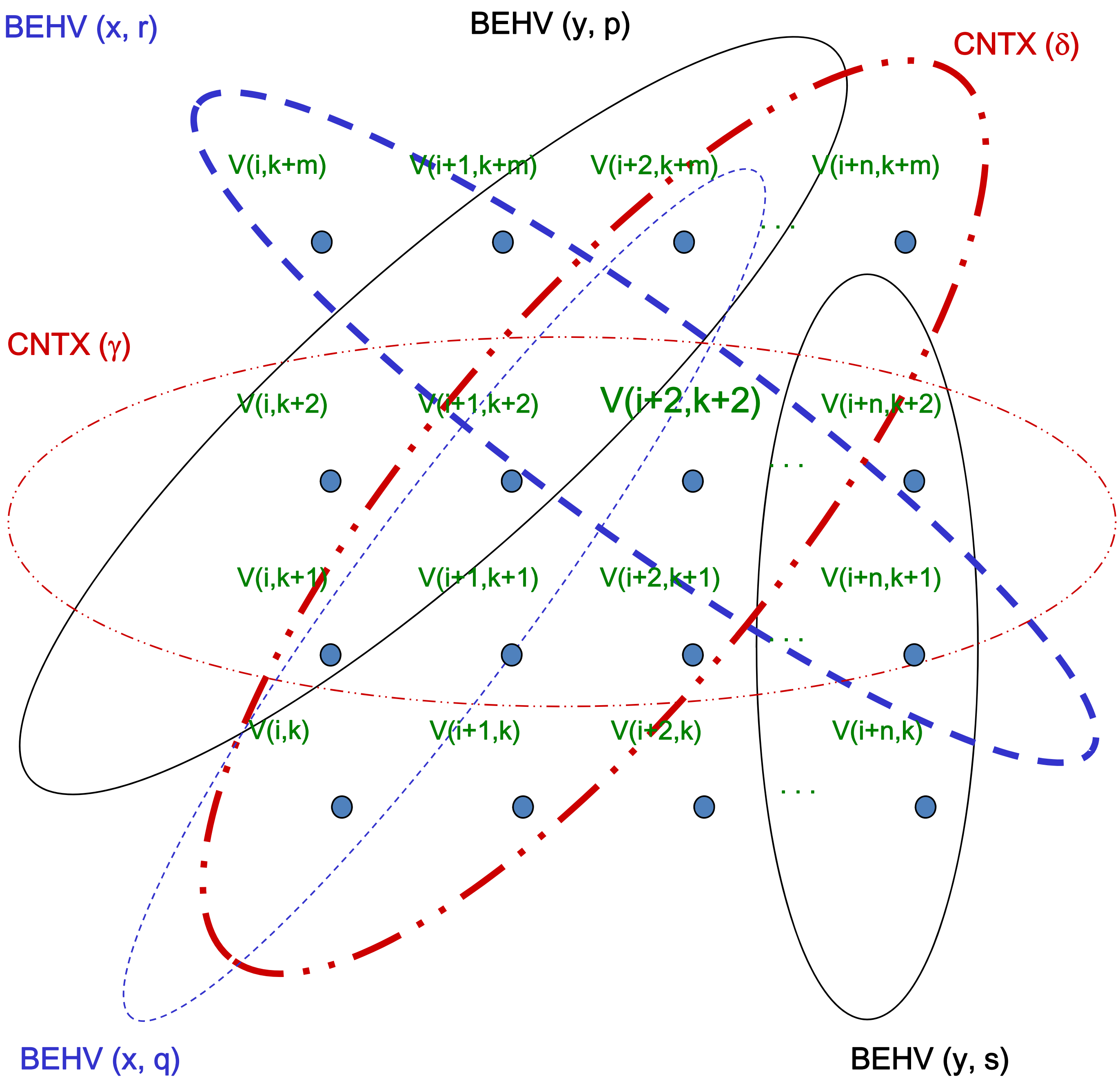
AU	Signification	No.	Correct	False	Missed	Confused	Recognition Rate
0	Neutral	20	17	3	0	0	85.0%
1	Inner Brow Raiser	10	8	0	2	0	80.0%
2	Outer Brow Raiser	27	20	1	5	1	74.0%
4	Brow Lowerer	24	21	1	1	1	87.5%
12	Lip Corner Puller	17	13	0	4	0	76.4%
15	Lip Corner Depressor	13	10	1	2	0	76.9%
26	Jaw Drop	24	14	0	10	0	58.3%
Total		135	103	6	24	2	76.2%
False Alarm: 4.4%, Missed: 17.7%							

Research Work Accomplished

- Working on the **development of clinical care grade solutions for information gathering from complex multi-sensor systems, under [M2.12] and [M2.13a],** we investigated JDL-type multisensor system architectures and intelligent data fusion algorithms able to incorporate a multitude of time- and location-dependent sensor-data, from which will be possible to extract relevant information about patient condition, clinical staff and activities, medication, supplies and equipment status, and working context, [Cretu2009], [Petriu2009], [Pozna 2009].

Heterogeneous network of “machine” sensor agents, symbiotic sensor agents, human sensor agents, *and intelligent sensor agents capable to comprehend human and animal behaviour (the human and animal, being used as transducers).*





Context-based plausible meaning of the specific behaviour of a human agent*:** Estimating the value **V** of an environmental parameter of interest based on the specific behaviour ***BEHV of a human agent, which is function of the respective parameter and the context ***CNTX.***

In the previous figure: the human agent "**x**" exhibits the *behaviour* **BEHV (x, r)** , which may occur for any of the following *environmental parameter values* **{V(i, k+m), V(i+1, k+m), V(i+2, k+2), V(i+n, k)}**, *in the context* **CNTX (d)** defined by the following values of the environmental parameter of interest **{V(i+2, k+m), V(i+n, k+m), V(i+1, k+2), V(i+2, k+2), V(i, k+1, V(i+2, k+1), V(i, k), V(i+1, k)}.**
It can be concluded that this **specific behaviour in the given context** occurred because of the **specific value V(i+2, k+2) of the environmental parameter** of interest, which is the value that is shared by the definition domains of the behaviour **BEHV (x, r),** and the context **CNTX (d).**

- Human sensor information is "fuzzy quantizied" while the machine sensor information, both the symbiotic analog_ transducer & human, and the fully automated digital one, is "sharp & concatenated quantized"

- It is possible to reduce the uncertainty of the measurements involving humans as sensors part of multisensor systems, by using Fuzzy Cognitive Maps, NNs, and Associative Memories.

- Dempster-Shafer theory of evidence approach is used to incorporate human-like uncertainty management and inference mechanisms in our context-aware multi-sensor data fusion system. This approach allows us to incorporate time-variable weights representative of sensor precision which will improve the sensor fusion accuracy in dynamic environments.

- Linguistic pattern recognition techniques and semantic model representations are used to develop a semantic level situation assessment system that will allow understanding of the dynamics of a complex scene based on multimodal sensor data streams.

Publications related to our work undertaken as a direct result of the hSITE technical agenda, but funded from complementary sources

[Chen2009a] Q. Chen, M.D. Cordea, E.M. Petriu, A.R. Varkonyi-Koczy, T.E. Whalen, “Human–Computer Interaction for Smart Environment Applications Using Hand Gestures and Facial Expressions,” *Int. J. Advanced Media and Communication, Vol. 3, Nos. 1/2,* pp. 95-109, 2009.

[Chen2009b] Q. Chen, F. Malric, Y. Zhang, M. Abid, A. Cordeiro, E.M. Petriu, N.D. Georganas, "Interacting with Digital Signage Using Hand Gestures", *Proc. ICIAR 2009, Int. Conf. Image Analysis and Recognition, (M. Kamel and A. Campilho - Eds), Lecture Notesin Computer Science Vol. LNCS 5627, pp. 347-358, Springer, Berlin, 2009.*

[Cretu2009] A.-M. Cretu. P. Payeur, E.M. Petriu, “Selective Range Data Acquisition Driven by Neural-Gas Networks,” *IEEE Trans. Instrum. Meas., vol. 58, no. 8, pp. 2634 -2642,* 2009.

[Petriu2009] E.M. Petriu, "Context-Aware Multi-Sensor Systems for Mission Critical Healthcare Applications," *IEEE Workshop on Self Monitoring, Analysis, Record and Testing for Adverse Reactions to Medication - SMART ARM 2009, IEEE Ottawa Section, in* Ottawa, ON, on Nov. 2009.

[Pozna2009] C. Pozna, R.-E. Precup, S. Preitl, E. M. Petriu, J. K. Tar, “Structure for Behaviourist Representation of Knowledge,” *Proc. 10th Int. Symposium of Hungarian Researchers in Computational Intelligence and Informatics - CINTI 2009, pp. 55-68,* Budapest, Hungary, Nov. 2009.

[Xhao2010] Y. Zhao, N.D. Georganas, E.M. Petriu, “Applying Contrast-limited Adaptive Histogram Equalization and Integral Projection for Facial Feature Enhancement and Detection,” *Proc. I²MTC 2010 – IEEE Int. Instrum. Meas. Technol. Conf.*, pp. 861-866, Austin, TX, USA, May 2010.