Sleep and Wake Functions Based On AIM Model for Multiple Vision System

Masahiko MIKAWA
Institute of Library and Information Science
University of Tsukuba
1-2 Kasuga, Tsukuba-shi, Ibaraki, JAPAN
Email: mikawa@slis.tsukuba.ac.jp

***Abstract*— The purpose of this paper is to design a mathematical model for a robot vision, that can express states of human's consciousness, and build a dynamical information processing system, that can change processing modes. Especially I take notice to a state transition between waking and sleep. A human being processes external information obtained from sensory organs and internal information accumulated within a brain flexibly and in parallel. The proposed mathematical consciousness model controls a ratio between external and internal information processing runs in parallel. In a waking mode, the external information is mainly processed. In a sleep mode, large proportion of processing is paused, or the internal information is mainly processed. The advantage of this system is that a resource of a computing machine is made effective use of. Because required processes are executed only when needed.**

I. Introduction

In recent years, a lot of real-time vision systems are put into practical use. In lots of vision systems, however, visual information is processed by series of procedures, that are selected and serialized previously depending on their purposes. It is difficult for these kinds of vision systems to be used in dynamical and complex environments. Moreover, the series of image processing are executed at full power regardless of use or non use. This wastes computer resources and electric power.

On the other hand, a human being has external and internal information processing systems that are flexible about dynamical and complex environments. Internal information means memories stored in a brain, and is used as knowledge for processing several external information. It is also processed when we dream. External information is obtained from external sensory organs. It is known that we obtain more than 80 % of external information from visual perception, and visual information is very important. The same is true of intelligent systems such as robot vision systems.

External and internal information processing systems of a human being are in parallel and multi-layered. Freud proposed that there were three levels of human mind, that are conscious, preconscious and unconsciousness[1]. He also described that the human mind is like an iceberg, it floats with one-seventh of its bulk hidden below water. The visible part of the iceberg means consciousness, and the unvisible part means preconsciousness and unconsciousness. This means that the unvisible part plays important and lots of roles. In a unconsciousness state, internal and external information processing modules are executed independent of human intentions in parallel. In a preconsciousness state, information is stored as memories, which can be easily summon into consciousness. In a consciousness state, sequence processes for processing particular information are executed intentionally in serial.

In the field of physiology, cycle models between rapid eye movement (REM) sleep and non-REM (NREM) sleep have been proposed[2][3], and these models are relative to the unconsciousness. In the vision research field, Marr and Hoffman proposed computational theories for understanding human beings' visual information processing systems[4][5]. Marr described that a process of visual perception has three stages of representations, that are the primal sketch, the $2\frac{1}{2}$ dimensional sketch and the 3 dimensional model representation. Although the primal and $2\frac{1}{2}$-D sketches are deeply related to the unconsciousness, Marr's theories mainly focus on the 3-D object recognition, which is related to the consciousness.

The purpose of this paper is to design a mathematical model for a robot vision system, that can express states of human's mind, and build a dynamical information processing system, that can change modes of processing smoothly. Especially I take notice to a state transition between waking and sleep. In a waking mode, external information obtained from a vision device are mainly processed. In a sleep mode, large proportion of external information processing is paused, or information accumulated in storage areas are mainly processed. The advantage of this system is that a resource of a computing machine is made effective use of. Because required processes are executed only when it is needed. There are many hypotheses of sleep mechanisms, for example, the restorative theories, the adaptive theories, the energy conservation theory and the programming-reprogramming theory. The idea of the sleeping model proposed in this paper is close to the passive theory of sleep[6]. According to this theory, that is an early idea in the field of sleep researches, sleep occurs to prevent fatigue or is caused by a lack of sensory stimulation.

Two kinds of sample applications executed in a stereo vision system reveal the validity and effectiveness of my proposed method. In one a monocular vision system is used. In the other a stereo vision system is used.

II. Hobson's AIM State-Space Model

Figure 1 shows the Active-Input-Modulation (AIM) state-space model[7][8][9] proposed by Hobson. He describes that states of human's consciousness can be expressed by levels of three elements, Activation, Input and Modulation of which the AIM model consists. Activation means an amount of information which is being processed. Input means a source of the information from either internal sources such as a memory area in a brain or external sources such as external sensory organs acquired information about an environment. Modulation switches external and internal processing modes. When an aminergic level is higher than a cholinergic level, external information is mainly processed. When a cholinergic level is higher than an aminergic level, internal information is mainly processed. For example, the states of waking, relax, REM sleep and NREM sleep are shown in Fig. 1. In a waking state, external information, that is obtained from external sensory organs is processed actively. In a REM sleep state, internal information accumulated in a storage area are processed actively. In a NREM sleep state, input and output gates are closed and processing power declines totally.

When a normal person falls asleep, at first the sleeper enters NREM sleep, which is a deep sleep. After the first NREM, he/she enters REM sleep and NREM sleep alternately until he/she awakes naturally or large stimuli interrupt the sleep.

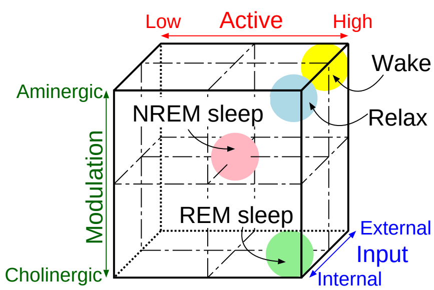

Fig. 1. Hobson's Spatial AIM model

III. Mathematical AIM Model

A. Mathematical Modeling for AIM

Figure 2 shows a relation among a mathematical AIM model proposed in this study and external and internal information processing systems. An external information processing system consists of plural external sensory organs, sampling parts, preprocessing parts and data processing parts. Raw data obtained by the sampling parts is processed by one or more preprocessors, and computed by each higher level data processor. It is important that external data is processed in real time and translates to meaningful information as soon as possible. An internal information system consists of internal data sources, data sampling parts and data processing parts. Information obtained from the external sensor, knowledge given previously, and so on, is stored in the internal data sources. In the internal information processing, following kinds of data are treated. One is a kind of data that it is difficult to process in real time, because it takes much time to process data. Another is what it is not necessary to process in real time. The third is a lot of data such as time-series data that must be processed all together. One external sensory organs is paired with one internal data source.

The AIM model dynamically controls ratios among the external and internal data samplers, preprocessors and information processors. The AIM model consists of elements S, A, I and M. The element S converts internal and external information into stimuli. The element A decides a frequency of each data processing. The element I decides values of parameters, that are used when stimuli are calculated in the element S. The element M decides a frequency of each data sampling. Each element consists of two sub-elements that have subscripts ex and in respectively. Here, the subscripts ex and in mean that the parameters relate to the external and internal information processing, respectively. The number of the sub-elements related to S, A and I is equal to the number of the preprocessors. The number of the sub-elements related to M is equal to the number of the sensor organs and the data sources.

Figure 3 shows an example of the variations of the sub-elements a_ex, a_in, i_ex, i_in, m_ex and m_in with time, that are included in a pair of the external and internal information processing flows. When the external sensory organ receives stimuli larger than a threshold ($s\_ex \geq th_{s\_ex}$ (threshold value)), the level of elements related to the external information processing is higher than that related to the internal ones. The level of external stimulus becomes lower than the threshold ($s\_ex < th_{s\_ex}$) at the time t_0. After the lapse of a constant period T_w, then the state is shifted from the waking to the relaxing. The relaxing state is kept while the external stimuli obtained by the other external sensor organs are larger than the threshold ($s'\_ex > th'_{s\_ex}$). All the levels of external stimuli become lower than the threshold ($s\_ex < th_{s\_ex}$ and $s'\_ex < th'_{s\_ex}$) at the time t_3. After the lapse of a constant period T_a, all levels of the elements become lower then the state is shifted from the relax mode to the NREM sleep. Moreover after the lapse of a constant period T_n, the values of each element increase and decrease periodically at frequency f_r. Where, the amplitude of vibration of each element related to internal processing is higher than that of each element related to external processing. Once again s_ex becomes larger than th_{s_ex} at the time t_7, the system state returns to the same state in $t < t_0$.

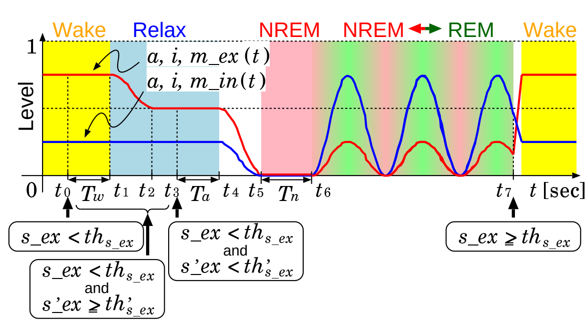

Fig. 3. Variations of a_ex, a_in, i_ex, i_in, m_ex and m_in with time

$a\_ex(t)$ and $a\_in(t)$ are given by the following equations. As it is able to express each the parameter related to the elements

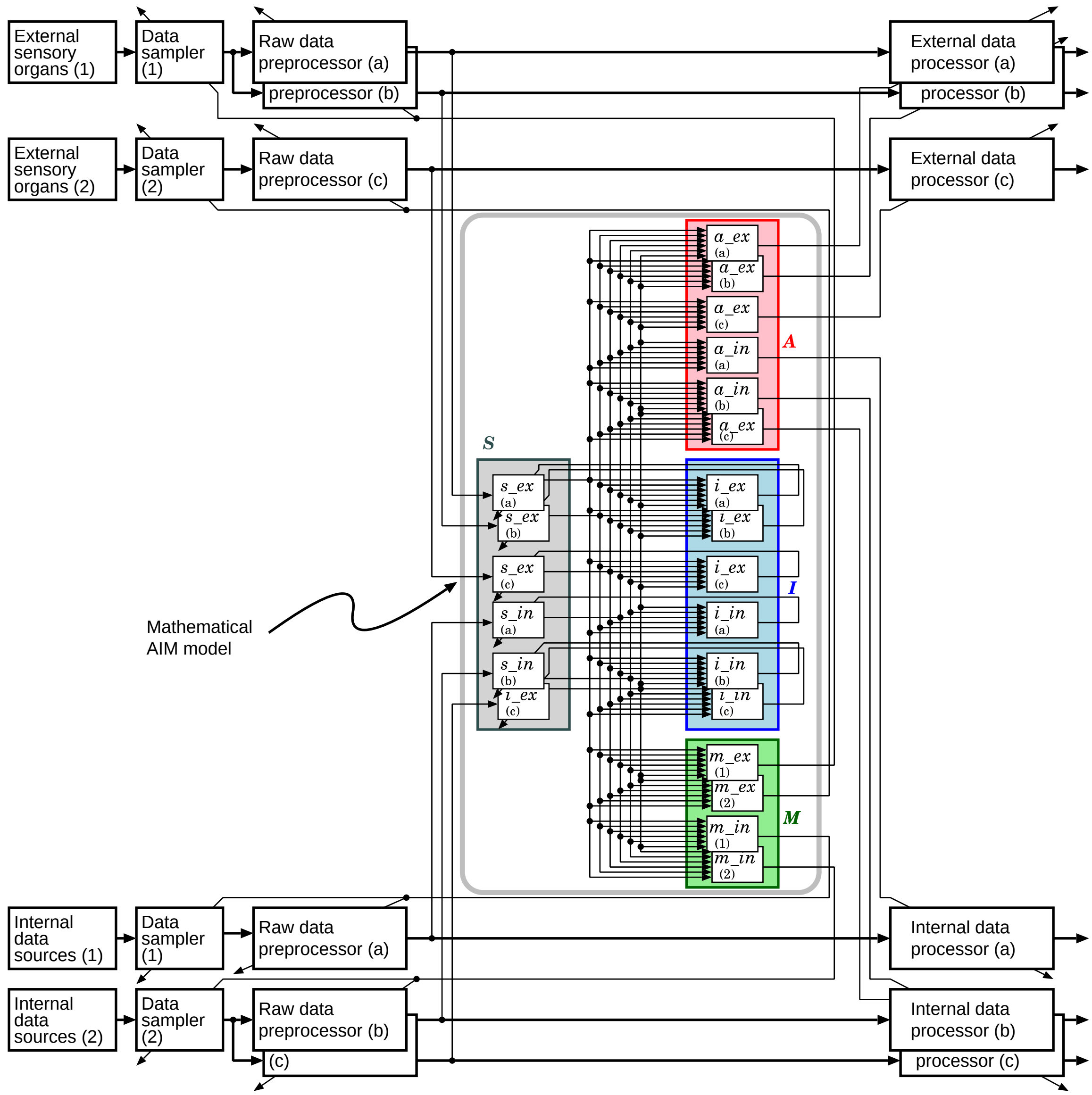

Fig. 2. Mathematical AIM model and information systems with plural external sensors

I and M in the same way, their descriptions are omitted here.

$$a\_ex(t) = \begin{cases} L_{w,a\_ex} + b_{a\_ex} & (t < t_1) \\ \zeta(L_{w,a\_ex}, L_{a,a\_ex}, t_1, f_{w,a\_ex}) & (t_1 \le t < t_2) \\ L_{a,a\_ex} + b_{a\_ex} & (t_2 \le t < t_4) \\ \zeta(L_{a,a\_ex}, L_{n,a\_ex}, t_3, f_{a,a\_ex}) & (t_4 \le t < t_5) \\ L_{n,a\_ex} + b_{a\_ex} & (t_5 \le t < t_6) \\ \xi(o_{r,a\_ex}, t_3, f_{r,a\_ex}) + L_{n,a\_ex} + b_{a\_ex} & (t \ge t_6) \end{cases} \tag{1}$$

$$a\_in(t) = \begin{cases} L_{w,a\_in} + b_{a\_in} & (t < t_4) \\ \zeta(L_{w,a\_in}, L_{n,a\_in}, t_1, f_{w,a\_in}) + b_{a\_in} & (t_4 \le t < t_5) \\ L_n + b & (t_5 \le t < t_6) \\ \xi(o_{r,a\_in}, t_3, f_{r,a\_in}) + L_{n,a\_in} + b_{a\_in} & (t \ge t_6) \end{cases} \tag{2}$$

where,

$$\zeta(\alpha, \beta, \tau_\zeta, f_\zeta) = \frac{\alpha - \beta}{2}\left(1 + \cos(\frac{2\pi(t - \tau_\zeta)}{f_\zeta})\right) + \beta + b \tag{3}$$

$$\xi(o, \tau_\xi, f_\xi) = o\left(1 + \cos(\frac{2\pi(t - \tau_\xi)}{f_\xi})\right) \tag{4}$$

The parameters $T_{w,a\_ex/in}$, $T_{a,a\_ex/in}$, $T_{n,a\_ex/in}$, $L_{w,a\_ex/in}$, $L_{a,a\_ex/in}$, $L_{n,a\_ex/in}$, $f_{w,a\_ex/in}$, $f_{a,a\_ex/in}$, $f_{r,a\_ex/in}$, $o_{r,a\_ex/in}$ and $b_{a\_ex/in}$, that are included in Eqs.(1) and (2), are independent of parameters included in other sub-elements such as i_en, i_in, m_en and m_in. Besides values of the parameters are designed depending on the purpose of an application and kinds of external and internal information processing.

The elements A, I and M are shown by the following equations using the sub-elements.

$$A(t) = \frac{1}{n+o}\left(\sum_{j=1}^{n} a\_ex_j(t) + \sum_{k=1}^{o} a\_in_k(t)\right) \quad (5)$$

$$I(t) = \sum_{j=1}^{n} i\_ex_j(t) - \sum_{k=1}^{o} i\_in_k(t) \quad (6)$$

$$M(t) = \sum_{l=1}^{q} m\_ex_l(t) - \sum_{m=1}^{r} m\_in_m(t) \quad (7)$$

where, n, o, q and r are the numbers of the external data processors, the internal data processors, the external sensory organs and the internal data sources, respectively. The AIM state-space model as shown in Fig. 1 can be expressed by Eqs.(5)∼(7). An example of variations of elements A, I and M with time is shown in Fig. 4. This example is resemble to a sleep pattern of normal person. The values of each parameter included in Eqs.(1) and (2) is defined as follows.

$$L_{w,a\_ex} = o_{r,a\_in} = 0.75 \quad (8)$$

$$L_{a,a\_ex} = 0.5 \quad (9)$$

$$L_{w,a\_in} = L_{a,a\_ex} = o_{r,a\_ex} = 0.25 \quad (10)$$

$$L_{n,a\_ex} = L_{n,a\_in} = 0 \quad (11)$$

$$b_{a\_ex} = b_{a\_in} = 0 \quad (12)$$

As it is able to express each the parameter related to the elements I and M in the same way, their descriptions are omitted here.

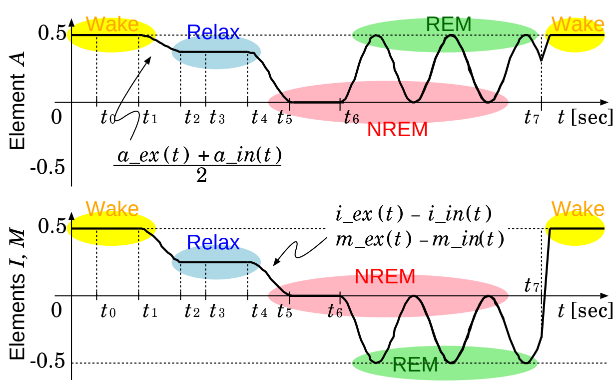

Fig. 4. Variations of elements A, I and M with time

In the following subsection, I describe a control of the internal and external information processing flows based on these behaviors of the elements A, I and M, functions of the element S and relations among S and the other elements.

B. Dynamic Control of Information Processing System Based On Mathematical AIM Model

External information obtained from the external sensor is converted to stimulation by the function of the element S. As shown in Fig. 2, the values of the sub-elements included in the elements A, I and M are determined based on the stimulus information. The operational states of the external/internal data samplers and processors are controlled dynamically by the functions of these elements included in the AIM model.

The frequencies of the external and internal information processing are determined by each level of a_ex and a_in. The frequencies of the external and internal processing increase or decrease in proportion to each level of a_ex and a_in, respectively. For example, when the level of a_ex becomes higher, the frequency of the external information processing increases. Then the external information processor acquires ability to calculate more volume of information. At that time, as shown in Fig. 3, the level of a_ex becomes lower, the frequency of the internal information processing increases. In the same way, the frequencies of the external and internal data samplings are determined based on the element M.

The element I decides the thresholds th_{s_ex} and th_{s_in} and sensor densities in order for the element S to recognize variations of information obtained from the external sensor or the internal source as stimuli. Here, the sensor density means quantity of points or sizes of areas for observing variations of internal or external information. The thresholds and the sensor densities vary in proportion to each level of i_ex and i_in. For example, when the level of i_ex becomes higher, the value of th_{s_ex} becomes lower and the sensor density becomes higher. This means that the system can perceive smaller variation of external and internal information more sensitively.

The state variations of the external and internal information processing systems with time are shown in Fig. 5. The graph in the top shows the frequency variations of the external/internal data processing with time. The graph in the middle shows the variations of both the threshold and the sensor density with time. The graph in the bottom shows the frequency variations of the external/internal data samplings with time. Where, all the conditions at the time $t_0 \sim t_7$ shown in Fig. 5 are the same with the conditions in Figs. 3 and 4.

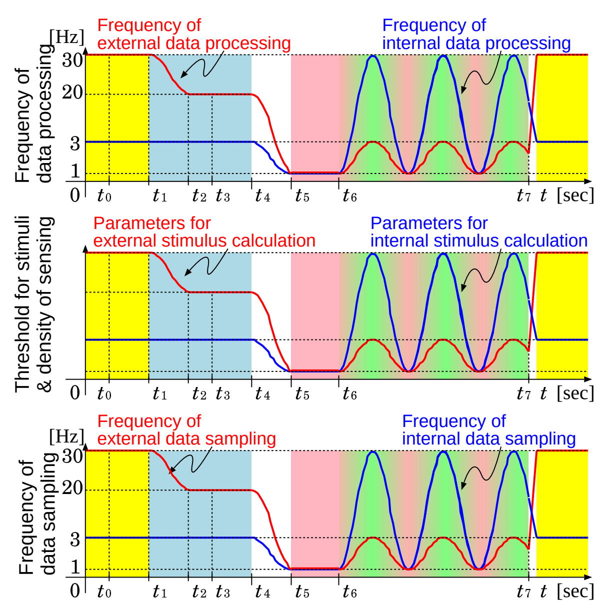

Fig. 5. Variations of each parameter included in internal and external information processing system with time

IV. Experimental Results

I made two sample programs to evaluate a validity of the proposed AIM model. One was a program using a monocular vision system. The other was one using a stereo vision system.

A. Experimental System

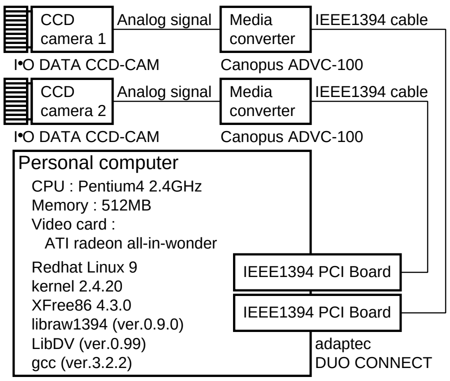

Fig. 6. Configuration of experimental system

Figure 6 shows a configuration of the experimental vision system. This system consists of two CCD cameras with an NTSC output, two media converters that can convert analog audio and video signals to digital audio and video (DV) signals, two IEEE 1394 interface PCI boards and a personal computer installed the Linux OS. The versions of the Linux kernel and the other software are also shown in Fig. 6.

I used the application programming interface (API)[10] developed by myself for capturing DV image. By using this API, video capturing processes, DV signal decoding processes, image displaying processes and processes programmed by a user can be multi-threaded and executed in parallel. It also is able to capture images from plural video devices in real-time (30 frames/sec). A captured image is represented in the YUV color space and composed of 720×480 pixels with 8 bits per each color. This API is suit to develop programs where plural external and internal processes are running in parallel.

B. Application Using Monocular Vision System

This program had the only external information processing flow and was simple. Moreover, only one CCD camera was used in the program. Brightness changes, which can be obtained by comparing each pixel included in two time series image frames, were used as stimuli.

Figure 7 shows a screen shot of the program. The state of the program was in the NREM sleep at this moment. The resolution of the display was 1024 × 768 pixels. The points where it was determined whether brightness change was more than the threshold th_{s_ex} or less were shown in green. In this program, the sensor density means the distance among green points. As its state was in the NREM sleep, the sensor density for the external information process was lowest. Moreover, as th_{s_ex} increased, the sensitivity to stimuli decreased. When the state is in the waking, all the points in the image function as sensors for perceiving stimuli. The whole image area becomes green in this state. Moreover, as th_{s_ex} decreases, the sensitivity to stimuli increases.

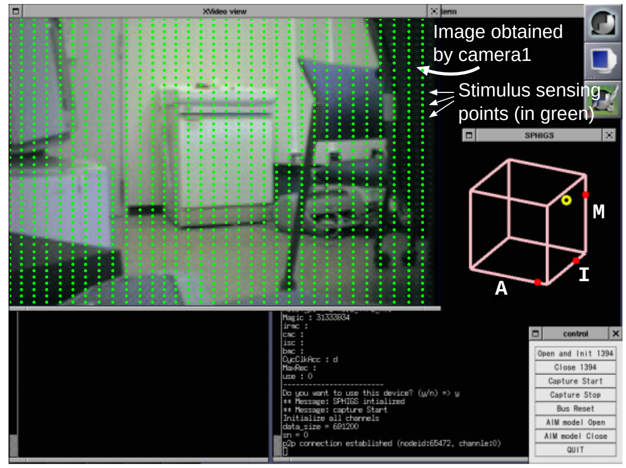

Fig. 7. Screen shot of sample application using monocular vision system

Figure 8 shows motion sequences and running time of each thread in this program. This program consists of a video capturing thread, a DV signal decoding thread, an image displaying thread and a thread for calculating the element A, I, M and S. The video capturing thread runs in a constant sampling time 1/30 sec. The capturing thread and the others are executed in parallel. The element A, I, M and S are calculated based on the brightness changes between two time-series decoded images. The operation frequencies of the capturing and decoding threads are decided by the value of a_ex. The threshold th_{s_ex} and the distance among the green points are decided by the value of i_ex. The operation frequencies of the displaying threads are decided by the value of m_ex. Although the sub-elements a_in, i_in and m_in were calculated, they were not used in this sample program.

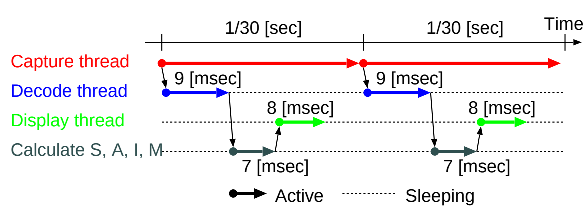

Fig. 8. Processing times of each thread

CPU utilization was measured while the program was running. Its variation with time is shown in Fig. 9. The values of each parameter included in Eqs.(1) and (2) were used what were shown by Eqs.(8)∼(12) and the follows equations. Where, only parameters related to the element A is written down in the same with Eqs.(8)∼(12).

$$T_{w,a\_ex} = T_{a,a\_ex} = T_{n,a\_ex} = 5 \text{ [sec]} \quad (13)$$

$$T_{w,a\_in} = T_{a,a\_in} = T_{n,a\_in} = 5 \text{ [sec]} \quad (14)$$

$$f_{w,a\_ex} = f_{r,a\_ex} = f_{w,a\_in} = f_{r,a\_in} = 20 \text{ [sec]} \quad (15)$$

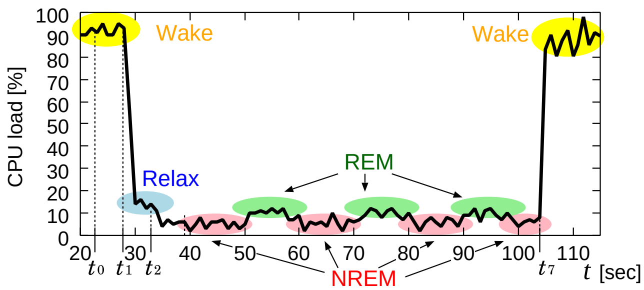

Fig. 9. Variation of CPU load with time using monocular vision system

In the waking state, image data was processed in 30 frames/sec, as a result, the CPU utilization was more than 90 %. After the change of image, in other words, stimuli from external information had disappeared at the time t_0, the CPU utilization fell around 15 % in the relaxing. After another T_n sec, it fell below 10 % in the NREM sleep. In the REM sleep, it became around 15 %. When the change of image occurred at the time t_7, the external information processing began to be executed actively again. The function of the AIM model makes it possible to make a margin of the CPU power for other processing while no external stimulus is present.

C. Application Using Stereo Vision System

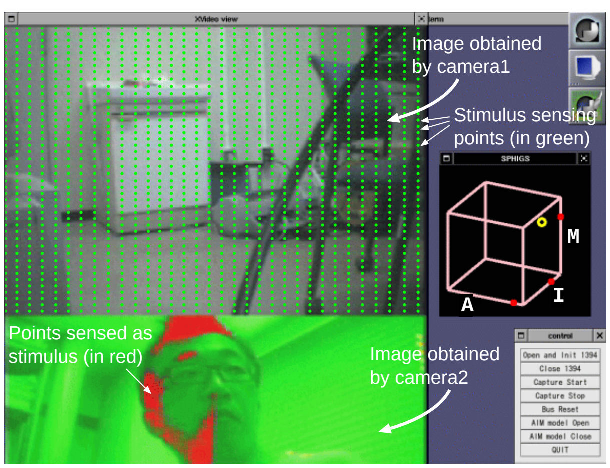

Fig. 10. Screen shot of sample application using stereo vision system

Figure 10 shows a screen shot of another sample program using a stereo vision system. This program consists of two external information processing systems. The each system is equipped with one CCD camera. The state of Fig. 10 was that the information processing system captured the upper image was in the relaxing mode, the other information processing system captured the lower image was in the waking mode. In the lower image, the points perceive brightness changes as stimuli were shown in red.

The variation of the CPU utilization with time is shown in Fig. 11. At the time t_0 stimuli related to the camera 1 disappeared, the camera 2 continued to receive stimuli. After T_n sec, the state of the information processing system related to the camera 1 shifted to the relaxing, and that related to the camera 2 stayed in the waking. As a result, the total CPU utilization fell below 80 %. When both of the cameras received no stimulus, the utilization decreased further. This result means that the function of the AIM model works well and can control the ratio between two information processing systems dynamically.

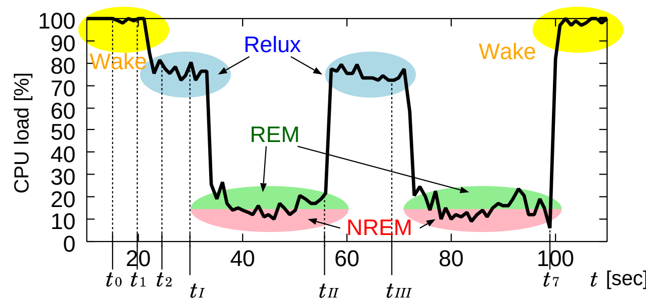

Fig. 11. Variation of CPU load with time using a binocular vision system

V. Conclusion

I proposed a new mathematical model based on the AIM state-space model for multiple vision. When a vision sensor receives stimuli, external information obtained from the visual sensor are mainly processed. When the external sensor hardly receives stimuli, an execution frequency of external information processing decreases, then large part of processing is paused. This means that required processes are executed only when needed, resources of a computing machine are made effective use of. I confirmed the validity and effectiveness of the proposed AIM model with experimental results.

In this paper, only very simple stimuli, that were brightness changes of pixels, were used in the application software. In future work, I will try to increase kinds of stimuli and implement internal information processes in the proposed system.

References

[1] Mark Solms. Freud returns. *Scientific American*, 290(5):56–63, 2004.

[2] J. Allan Hobson, R. Lydic, and H. Baghdoyan. Evolving concepts of sleep cycle generation: From brain centers to neuronal populations. *Behavioral and Brain Sciences*, 9(3):371–448, 1986.

[3] Robert W. McCarley, Robert W. Greene, Donald Rainnie, and Chiara M. Portas. Brainstem neuromodulation and REM sleep. *Seminars in The Neurosciences*, 7(5):341–354, 1995.

[4] David Marr. *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman and Company, 1982.

[5] Donald D. Hoffman. *Visual Intelligence : How We Create What We See*. W. W. Norton & Company, 1998.

[6] Michel Jouvet. Biogenic amines and the states of sleep. *Science*, 163(862):32–41, 1969.

[7] J. Allan Hobson, Edward F. Pace-Schott, and Robert Stickgold. Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, 23(6):793–842, 2000.

[8] J. Allan Hobson. *Consciousness*. W. H. Freeman & Co., 2001.

[9] J. Allan Hobson. *The Dream Drugstore*. The MIT Press, 2001.

[10] Masahiko MIKAWA. A programming interface for IEEE 1394 video device in robot vision. In *SICE Annual Conference 2003 in Fukui : Proceedings*, pages 2132—2137, 2003.