International Journal of Biological, Life and Agricultural Sciences
ISSN: 2415-6612
Vol:8, No:7, 2014

Various Advanced Statistical Analyses of Index Values Extracted from Outdoor Agricultural Workers Motion Data

Shinji Kawakura, Ryosuke Shibasaki

***Abstract*—**We have been grouping and developing various kinds of practical, promising sensing applied systems concerning agricultural advancement and technical tradition (guidance). These include advanced devices to secure real-time data related to worker motion, and we analyze by methods of various advanced statistics and human dynamics (e.g. primary component analysis, Ward system based cluster analysis, and mapping). What is more, we have been considering worker daily health and safety issues. Targeted fields are mainly common farms, meadows, and gardens. After then, we observed and discussed time-line style, changing data. And, we made some suggestions. The entire plan makes it possible to improve both the aforementioned applied systems and farms.

***Keywords*—**Advanced statistical analysis, wearable sensing system, tradition of skill, supporting for workers, detecting crisis.

I. Introduction

THE latest IT and statistical techniques have been performing interventions in actual agricultural fields. The combinations in themselves are thought as rather beneficial. According to “Japanese society of Agricultural Informatics”, a critical shortage of young, beginning agricultural workers and the ubiquity of methods of outdated traditional skills of mature (experienced) workers to immature (inexperienced, novice) [1]-[4].

In past studies, however, there is no research concerning “concrete suggestions for real users” to improve their physical motions of agricultural work, especially for beginners and immature workers. Therefore, our research supports such agricultural workers (and their daily actions) by utilizing electronic technologies, human dynamics, and advanced various statistical methods (e.g. using major indicators of statistics and those of wave-analysis). Providing workers with concrete feedback (e.g. by oral, visual, and some pieces of data-paper) is inevitable regardless to the period.

For the future, several sensing techniques and existing large-size farming machinery, the newest IT services and other tools should be integrated, and uniformly combined. Under these circumstances, we have been building integrated systems to raise productivity and the quality of skills’ tradition in real agricultural fields by analyzing records and extracting useful insights and clues for the future appropriately. But also, we suppose that statistical analyses concerning “flesh-and-blood” workers movements are meaningful significant.

Shinji Kawakura is with Center for Spatial Information Science, The Graduate School of University of Tokyo, Chiba, Japan (phone: +81-47-136-4307; e-mail: kawakura@iis.u-tokyo.ac.jp).

Ryosuke Shibasaki is with the Department of Socio-Cultural and Socio-Physical Environmental Studies, The University of Tokyo, Chiba, Japan (phone: +81-47-136-4307; e-mail: shiba@csis.u-tokyo.ac.jp).

II. Method

A. Subjects

We selected 15 subjects (see Table I). We gathered “hard-to-find” mature (experienced) and immature workers who do not have remarkable (serious) character both mentally and physically. The main criteria of not having any serious diseases, remarkable habits, and specific careers (especially in sports and martial arts) were developed from preliminary hearing and observing basic time-line data. Mature worker S.D. (Standard Deviation) values concerning each index are rather higher than those of immature workers.

TABLE I
Data Concerning Subjects

Index	Mature (Experienced) N = 7			Immature (Inexperienced, novice) N = 8		
	range	Ave	S.D.	range	Ave	S.D.
Age (year)	31 - 74	62.52	14.2	23 - 34	5.6	3.58
Experience (year)	2 - 60	34	18.1	0	0	0
Stature (cm)	155 - 173	164	5.5	170 - 180	174	3.2
Weight (kg)	55 - 85	70	8.9	58 - 78	67	7.5

B. Targeted Task

After conducting some interviews about subjects’ daily tasks in general farmlands (both outdoor and indoor), we made original categorization of workers common tasks (see Table II). And, we selected "Work in a semi-crouching position; Cultivating (digging up) by a hoe", the reason is that this movement is repetitive, with a rather large "full-body motion". Every trial was consisted of "30 swings (digging up)", and every subject performed three trials successively on the same day with inserting the few-minutes break. This style and number of sets seemed appropriate for this study, in light of past, this scientific field’s studies (see Fig. 1).

C. Electronic System

Firstly, we reviewed past studies and accomplishments including industrial goods, and patents [5]-[7]. And, we discussed our findings with researchers in Agricultural Informatics, actual workers, and farmland managers.

TABLE II
CATEGORIZATION OF TASKS

Categorization	Description
Work in a semi-crouching position	Cultivating (digging) by a hoe
Work in a sitting position	Harvesting from general root plants by hand, cumbersome; successive procedures both in any warehouses and outer fields (e.g. digging holes and overlaying clay, spraying water)
Work in a standing position	Cutting branches by shears and saw
Work alternating in a sitting and a standing position	Suspending onions or garlic on beams in an outhouse

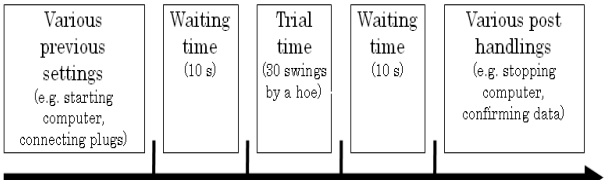


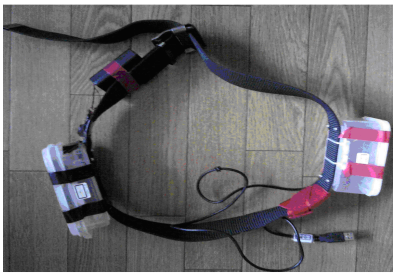
Fig. 1 Timeline of each trial

What became evident was that past studies and politics prove insufficient when considering such actual, realistic problems. For instance, human-dynamics oriented (particularly using various acceleration data), practical studies using attached Wearable Sensing Systems (WSs) are almost non-exist.

In light of our results [5], we shaped plans, and designed system constructions to measure and to analyze acceleration data (and for reference, angular velocity data). In addition, we have set the experimental conditions (e.g. size and color of wear, weather, hoe-related factors, the position of standing and grasping a hoe, posture of the starting position) uniformly as much as possible. We have provided directions to the 15 subjects and their managers simultaneously. And, we have been explaining to subjects that the main goals are to improve and to upgrade their skills and safety, including the easiness of their sophisticated skills' tradition. Considering the aforementioned factors, we made approximate schedules and designed each part (e.g. mechanical constructions, computer systems, and kinds of processes).

After such designing, we searched and selected some existing, promising techniques and modules, and constructed original integrated electronic systems (see Figs. 2, 3). Considering them, we have been doing some indoor operational testing to estimate system utility and suitability, in advance of outdoor experiments. We obtained a large amount of basic, CSV-shaped time series data of acceleration (and angular velocity data) concerning each part of user body. We also performed analyzing to certify their validity by taking basic kinds of postures and adding 3-axis vibrations.

As reference data, time series of data related to subjects' various daily, typical postures, and typical actions (e.g. standing upright, sitting on chair, lying down, walking, running, etc.) were also obtained. After them, we analyzed those data by original programs using Visual Basic 2010, Visual C++2010.



(a)

(b)

Fig. 2 Original vest-shaped and belt-shaped WSs, including Microcomputers and various devices (e.g. store-bought sensors)

(a)

(b)

Fig. 3 Hoe connected to sensors and laptop PC, one subject equipped with measuring modules and knapsack with laptop PC connecting to various modules

D.Survey Sheet and Statistical Analysis Software

We considered and selected major items (indicators) in the field of "statistics", "human dynamics" and "exercise physiology" to obtain the "features" and "changes" of agricultural workers (see Table III). In past studies, however, there is almost no such fusion in this field; we conducted after each outdoor trial set.

Some of them have been made by and used by Japan Association of Industrial Health and other health organizations, these are considered reliable. We asked about subjects' fitness habits, smoking habits, and their backache, just to make sure about subjects.

What is more, we chose two specific scales (The Visual Analogue Scale (VAS) and The Borg RPE Scale (RPE: Rating of Perceived Exertion)) to measure worker fatigue and feelings against strength of task. "VAS" is a psychometric response scale used in questionnaires. It is an instrument for measuring subjective characteristics or attitudes. When responding to a VAS item, respondents specify their level of agreement to a statement by indicating a position along a continuous line between two end-points. This continuous (or "analogue") aspect of the scale differentiates it from discrete scales. "RPE" measures perceived exertion, in sports and particularly in exercise testing. In medicine, this scale is used to document a patient's exertion during a test, and sports coaches use the scale to assess the intensity of training and competition. The original scale rated exertion on a scale of 6-20.

TABLE III
ITEMS IN SURVEY SHEET

Category	Index	Range of score (point)
Basic information	Name, affiliation, occupation, stature, weight, pre-existing disease	These depend on contents
Low back pain (LBP)	Experience of Low back pain	No experience of LBP (0), Experience LBP in the past (1), Now having LBP (2)
	Frequency in the present workplace	No (0), Sometimes (1), Frequently (2)
	Frequency in the past workplace	No (0), Yes (1)
Daily successive fatigue	Frequency of continuing fatigue from the previous day	No (0), Rarely (1), Sometimes (2), Always (3)
Drinking and smoking habit	Alcohol consumption	No (0), A few times a month or a year (1), Everyday or a few times a week (2)
	Tobacco consumption	Non smoker (0), Past smoker (1), Smoker (2)
Sport habit	During spare time	No (0) Yes (1)
	In the past	Non (0), A little in the past (1), Regularly in the past (2)
This trials' feeling of fatigue	Indicators in VAS (Visual Analogue Scale) and RPE (Borg RPE Scale) test, and oral, general question	VAS (0～100), RPE (6～20), and open-ended question
Usability of the systems	Load of the systems and the tasks, load of the work posture Fatigue of muscles	Five-grade evaluation (0～5) , and open-ended question

On the other hand, after basic trials in outer fields, we defined some major indicators, concerning vertical acceleration value of vertical direction: 1) Maximum value, 2) Minimum value, 3) S.D. (Standard Deviation), and 4) DC (Direct Current) component.

Since these studies are preliminary and tentative, we selected major indexes according to previous studies. Bao et al. [7] have used such indicators. "S.D." and "DC component" especially have been shown as useful values, considering discrimination concerning subjects and tasks. We probatively calculated and used large-sized "correlation matrix", "analysis of principal component". The reason is that these methods are considered rather comprehensive methods to categorize (group) catch-call persons and data from subjects.

As for calculating software, we selected Free-software "23_TAHN.xls" provided by Tomio Kikuchi in "KTS&C" and original "Add-in of Excel for multi-variable analysis" provided by "iStat, inc." in Japan. These were selected due to their utility and general-purpose properties, also considering our future prospects (testing other calculations, combining with other methods, and anticipating future needs).

III. RESULTS AND DISCUSSION

After frequently discussing, actual experiments (both indoor and outdoor), in light of various past results, we conducted not only necessary operating verifications of each function of the aforementioned systems' factors, but also essential experimental results (time series of data) are obtained concerning WSs in themselves. Each performance was observed as appropriate only for actions not containing "moving over" such as walking or cycling, targeted in this study.

The average value of VAS differed little (about 1.4 points of 100 degrees); on the other hand, RPE showed an apparent difference (about two points of 15 degrees). These results are believed to have been caused by differences in experience, skills, and physical fitness. Concerning their Low Back Pain (LBP), "continuing fatigue" making a reference was not easy. But mature "Frequency of LBP in the present workplace" was significantly higher than that of immature workers; it was typical Japanese trend and ideal for this study's one aim. Average values concerning aforementioned items (see Table III) in survey slips are described in Table IV. Concerning most factors, mature workers S.D. values are likely to be higher than immature workers.

Only concerning matrix of simple correlation coefficient (r) from the values of aforementioned survey slips' items and a hoe and waist time series of vertical acceleration data, there were many strong correlations (over 0.5 points and under -0.5 points of *r*). We remark only overly strong values of *r* (over 0.75 points and under -0.75 points) in this paper.

In light of results in Table V, for summing up, average values concerning" 3rd waist S.D.", "3rd hoe Min" are observed as significantly effective and critical factors, except for their age and experience. Next important factors are "1st hoe Max", "1st hoe DC", "3rd hoe DC", "3rd waist Min", and "3rd waist DC". We think, as Bao (2003) [7] et al. and other researchers have mentioned, these indexes are effective for grasping such

repetitive, dynamic actions. After them, the third priority group is "Past sports experience" and other factors.

Average values concerning time series of data after constant time (about 10 - 15 minutes) progress, especially concerning waist vertical acceleration the DC component could count on not only the most important. In other words, it is user trunk's "postures". But these postures seem to depend on worker age and experience. It is not easy to reform, but that is why we think these systems are helpful for it.

As we omit the details in this paper; weight, fatigues and movement intensity were little concerned with objective indexes extracted from vertical acceleration data noted in the foregoing. The changes in time-line styled data concerning acceleration data of mature workers are completely smoother than those changes of immature workers. Sudden movements are virtually absent, likely because their habituation and stability are due to years of experience. These characteristics could be observed visually too.

In light of results in Table VI, "minimum value" and "DC component" concerning hoe movement (hoe vertical acceleration), and whole waist movement (vertical acceleration) after a three trials (minutes) are thought as key factors.

That is why immature workers should carefully think about and pay attention to such factors when they are involving in such tasks.

TABLE IV
BASIC DATA CONCERNING SURVEY SLIPS' ITEM OBTAINED AFTER ONE SET OF TRIALS

Index	Mature			Immature		
	Range	Average	S.D.	Range	Average	S.D.
VAS	0 - 68.1	28.2	24.6	0 - 73.6	26.8	23.3
RPE	9 - 12	10.1	1.17	12 - 13	12.1	0.35
Experience of LBP	0 - 1	0.38	0.48	0 - 1	0.57	0.49
Frequency of LBP in the present workplace	0 - 1	0.50	0.50	0 - 1	0.29	0.45
Frequency of LBP in the past workplace	0	0	0	0 - 1	0.14	0.35
Frequency of continuing fatigue from the previous day	0 - 2	1.25	0.83	1 - 2	1.86	0.35
Alcohol consumption	0 - 2	1	0.87	0 - 2	1	0.53
Tobacco consumption	0 - 2	0.63	0.70	0 - 1	0.14	0.35
Sports habit during spare time	0 - 1	0.38	0.48	0 - 1	0.86	0.35
Sports habit in the past	0 - 1	0.38	0.48	1 - 2	1.57	0.49

TABLE V
SIMPLE CORRELATION COEFFICIENT (R)

Range	$-1.00 < r \leqq -0.75$	$0.75 \leqq r < 1.00$
Index	• age - stature • 3rd waist Max - experience • 3rd waist S.D. - age, experience, 3rd hoe Min, 3rd hoe DC, and 3rd waist Min	• age - experience, 3rd hoe Min, 3rd waist Min • 1st hoe Max - 1st hoe S.D., 3rd hoe Max, "Frequency of continuing fatigue from the previous day" • 1st hoe Min - 1st hoe DC, 3rd hoe Min • 1st hoe S.D. - 3rd hoe S.D. • 1st hoe DC - 3rd hoe DC • 3rd hoe Min - 3rd hoe DC • 1st waist Min - 1st waist DC, 3rd waist DC • 1st waist DC - 3rd waist DC • 3rd waist Max - Past sports experience • 3rd waist Min - 3rd waist DC • 3rd waist S.D. - Past sports experience

TABLE VI
CORRELATION RATIO (H^2: 0 - 1) BETWEEN MATURE AND IMMATURE SUBJECTS ($H^2 \geqq 0.50$, $0.50 > H^2 \geqq 0.25$)

Index				η^2			Index		η^2		Index	η^2
Age				0.75								
Experience				0.62								
Stature				0.56								
Weight				0.03								
VAS				0.0009								
RPE				0.56								
From the time series of acceleration data	1st trial	Hoe	Max	0.03	From the time series of acceleration data	3rd trial	Hoe	Min	0.56	From survey slips	Experience of LBP	0.04
			Min	0.39				S.D.	0.01		Frequency of LBP in the present workplace	0.05
			S.D.	0.01				DC	0.77		Frequency of LBP in the past workplace	0.05
			DC	0.61		1st trial	Waist	Max	0.32		Frequency of continuing fatigue from the previous day	0.18
	3rd trial	Hoe	Max	0.01				Min	0.25		Alcohol consumption	0
								S.D.	0.003		Tobacco consumption	0.16
								DC	0.10		Sports habit during spare time	0.24
						3rd trial	Waist	Max	0.47		Sports habit in the past	0.60
								Min	0.65			
								S.D.	0.75			
								DC	0.39			

Figs. 4 (a), 4 (b), 5 (a), and 5 (b) show changes in cluster analysis charts utilizing "Ward system" and standardized data by the "23_TAHN.xls" program.

The first chart includes the values of age, experience, stature, and weigh, Fig. 4 (b) is the chart added the points of VAS and RPE to Figs. 4 (a), 5 (a) chart was the indicators concerning hoe vertical acceleration to Figs. 4 (a), 5 (b) is added the indicators concerning low-back pain, daily habits (in Table III) to Fig. 4 (a).

We just could experimentally observe the gradual changing concerning both the shape of charts and the calculating steps' "square distance" (e.g. in the graph of Fig. 5 (a), the values are 72.3 ⇒ 17.0 ⇒ 7.61…). These chart forms and the allocations of subjects are likely to change flexibly with varying raw data. This method would certainly be useful, but we should discreetly select factors to be included into these data in such kinds of studies.

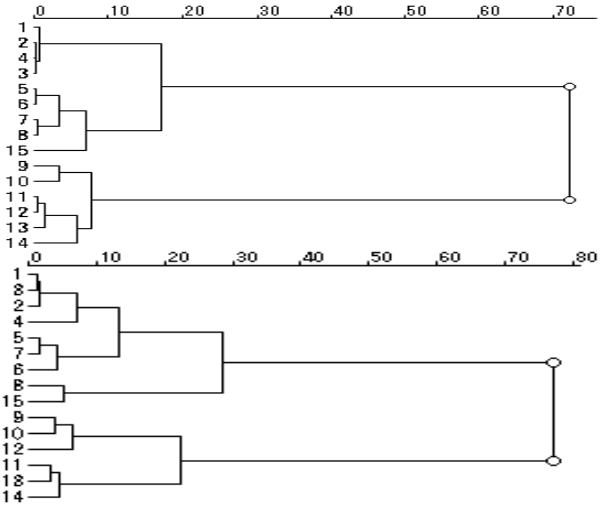


Fig. 4 (Upper: (a), under: (b)) Changes of cluster analysis' chart

On the other hand, we performed analysis of principal component by the aforementioned "Add-in of Excel for multi-variable analysis". One of the main targets is to diagnose "The (hidden) factors that affect motions and user condition". These output data are seen in Figs. 6-11.

Figs. 6, 7, and 8 include basic information and values concerning hoe vertical acceleration. Figs. 9, 10, and 11 include basic information and values concerning survey slips (see graph captions and descriptions).

Figs. 6, 7 suggest the primary ingredient as "The major (hidden) factor that affects the horizontality and the minimum value, concerning both hoe (hand) and waist acceleration due to age and experience"; and the secondary ingredient as "the factor affects altitude to fling up a hoe, hoe movement's volatility and horizontality". These are rather clear.

From Fig. 8, the primary ingredient's "influence rate" was thought stronger than the secondary one judging from this scatter gram.

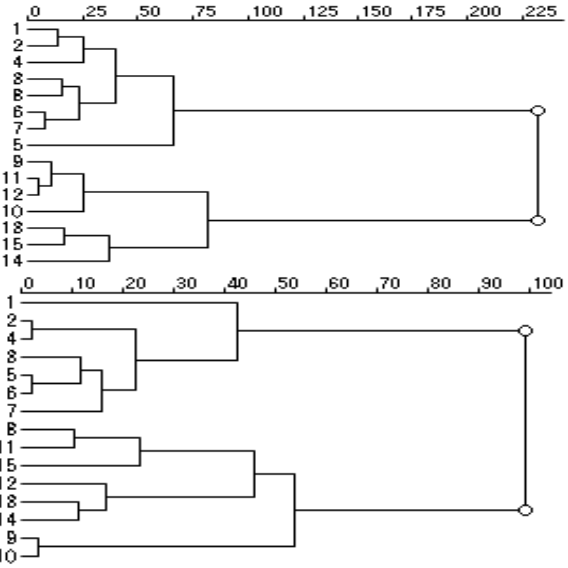


Fig. 5 (Upper: (a), under: (b)) Changes of cluster analysis' chart

About Figs. 9, 10, we could call "the primary ingredient as "The major (hidden) factor that affects the perceived exertion, recent LBP, remaining fatigue, and bad habits for health". The secondary ingredient could be labeled as "the factor affects perceived exercises' intensity, long-time LBP, and bad habits for health". These two ingredients are also intuitive, and daily unhealthy habits and their age certainly affect perceived exertion and intensity concerning this kind of agricultural work.

From Fig. 11, the primary ingredient's "intensity" was also thought stronger than the secondary one, but not so simple "scattering condition", so more consideration in needed.

At the same time, we observed that four ingredients may have been derived from characteristics of both mature and immature worker skills, their physical condition, and body (muscle) strength. Of course, both other conditions and examples must be considered more in after cases. And still, other problems also remain with these analyses, which may be solved by optimizing and regularizing various conditions (e.g. subject numbers, and variation) to the extent possible.

IV. Conclusion

Our studies have exemplified several basic steps of the aforementioned styles of systems and methods by examining actual work sites. Our research has obtained promising prospects, especially concerning the fusion of Agricultural Informatics, statistics, and Human Dynamics to some extent. But their endurance and validity must be confirmed later; above all, we should add a greater variety of worker data into these statistical data.

And, other recent methodologies of human dynamics and visual data analysis (e.g. higher mathematics) should be tested and incorporated. What is more, we have been obtaining more various data, and been sending instructions to actual workers, particularly about "skill tradition" and their daily tasks, after certifying their usefulness.

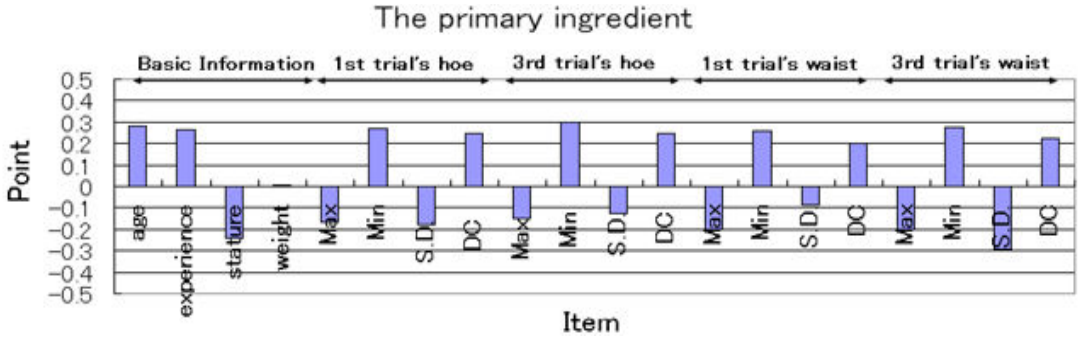


Fig. 6 Bar graphs concerning the primary ingredient of analysis of principal component using basic information and values concerning hoe vertical acceleration

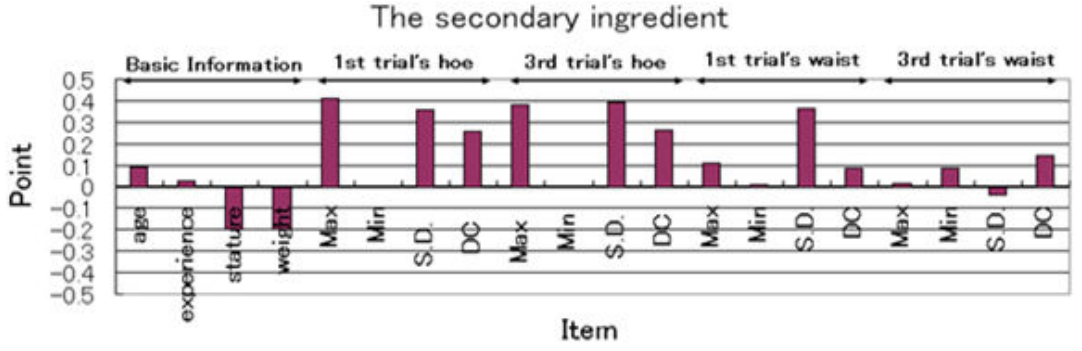


Fig. 7 Bar graphs concerning the secondary ingredient analysis of principal component using basic information and values concerning hoe vertical acceleration

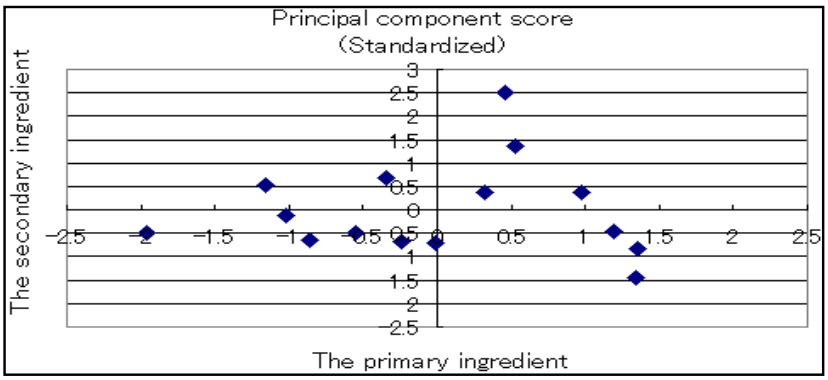


Fig. 8 Scatter gram concerning analysis of principal component using basic information and values concerning hoe vertical acceleration

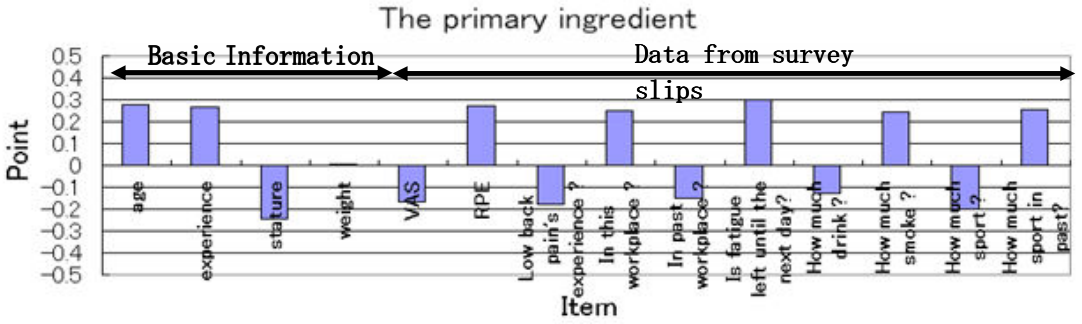


Fig. 9 Bar graphs concerning the primary ingredient of analysis of principal component using basic information and value concerning survey slips' items

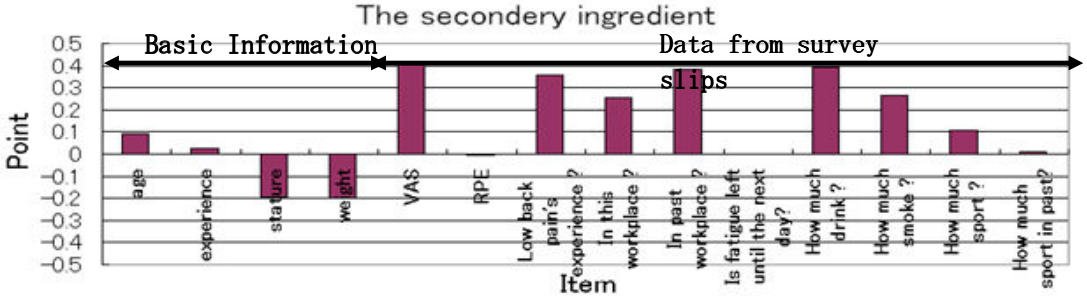


Fig. 10 Bar graphs concerning the secondary ingredient of analysis of principal component using basic information and value concerning survey slips' items

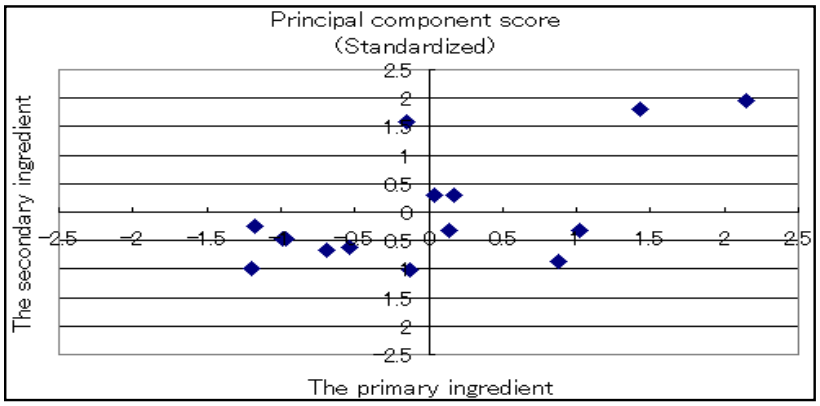


Fig. 11 Scatter gram concerning analysis of principal component using basic information and value concerning survey slips' items

We also think that, reviewing the entire results and work products will prove that the measure of precision for diagnosing critical situations (e.g. heavy diseases, injuries) will be improved. The stability of the whole system against sudden accidents will also be improved in the future. The reason is that irregular posture data can certainly be diagnosed.

In addition, we plan to involve other natural outdoor farms, artificial indoor farms, and other industries (situations) in considering robustness and stability. Furthermore, utilizing these results and many sets of experimental evidence, we plan to launch practical supporting projects for workers. And, from the view point of global agricultural dynamics, we have some plans to launch to other countries.

These trials have been certainly challenging so far; however, they will cover worker contribution to agricultural industries comprehensively.

ACKNOWLEDGMENT

My heartfelt appreciation goes to Dr. Tokihiro Fukatsu, Mr.. Takuji Kiura, Dr. Masayuki Hirafuji (NARO, NARC), Tsukuba University, CSIS and IIS at The University of Tokyo, and GRENE project in Japan, who have provided careful support, feedback, and valuable comments.

REFERENCES

[1] Y. Fujii, T. Nanseki, H. Kobayashi, and K. Nishitani, "Information management and cultivation of employee capabilities at large-scale paddy field farms: a case study of raising rice seedlings," Agricultural Information Research, vol. 21(3), pp. 51-64, 2012.

[2] Y. Fujii, T. Nanseki, H. Kobayashi, and T. Kojima, "The characteristics of expert know-how in agricultural planning on large-scale paddy field farms: a case study of a corporate farm in shiga prefecture," Agricultural Information Research, vol. 22(3), pp. 142-158, 2013.

[3] T. Nanseki, S. Takeuchi, and Y. Shinozaki, "Business development, ICT use, and personal training in agricultural corporations: an analysis of nationwide questionnaire survey," Agricultural Information Research, vol. 22(3), pp. 159-173, 2013.

[4] A. Shinjo and M. Kudo, "The practical use of IT in agriculture: the movement into high-value-added crops and integrated solutions," The Journal of the Institute of Electronics, Information, and Communication Engineers, vol. 96(4), pp. 280-285, 2013.

[5] S. Kawakura and R. Shibasaki, "Supporting systems for agricultural workers's skill and security," Proceedings of ACRS 2013, pp. 71-77, 2013.

[6] P. Mistry, P. Maes, and L. Chang, in Massachusetts Institute of Technology. WUW-wear Urworld: a wearable gestural interface. Available: http://dspace.mit.edu/handle/1721.1/61366, 2009.

[7] L. Bao, "physical activity recognition from acceleration data under semi-naturalistic conditions," unpublished master thesis, MIT, Boston, 2003.

Shinji Kawakura Center for Spatial Information Science, The Graduate School of University of Tokyo / Kashiwa-shi, Chiba, Japan. Date of birth: July 14, 1978. B.A in control system engineering, Tokyo Institute of Technology, 2003. M.A. in human-factors engineering, Tokyo Institute of Technology, 2005. (Meguro-ku, Tokyo, Japan).

Past Career: system engineering, researching in private companies.

Development and Verification of Wearable Sensing Systems with Real-time Spoken Commands for Outdoor Agricultural Workers with High-performance/low-cost Ambient Sensor Cloud System Using Open-Fs (Kawakura (Agricultural Information Research) and Dr. Shibasaki). Suggestions of Measuring and Analyzing Methods by Wearable Systems to Sense Outdoor Agricultural Worker Motion (Kawakura (Agricultural Information Research) and Dr. Shibasaki).

Ryosuke Shibasaki Department of Socio-Cultural and Socio-Physical Environmental Studies, The University of Tokyo / Kashiwa-shi, Chiba, Japan. Dr. Engineering. Professor of Center for Spatial Information Science, University of Tokyo.