Wafer Sort Bitmap Data Analysis Using the PCA-Based Approach for Yield Analysis and Optimization

Yeou-lang Hsieh, Gwo-hshiung Tzeng, *Fellow, IEEE,* T. R. Lin, and Hsiao-cheng Yu

***Abstract*—Yield analysis is one of the most important subjects in IC companies. During the initial stage of new process development, several factors can greatly impact the yield simultaneously. Traditionally, several learning cycle iterations are required to solve yield loss issues. This paper describes a novel way to diagnose yield loss issues in less iteration. First, the failure classification of bitmap data is transferred to a new basis using principal component analysis. Second, the defective rates are calculated and the original bitmap data is reconstructed in the principal basis, allowing the yield loss space to be generated by Cluster Analysis. Third, physical failure analysis samples can be selected to solve yield loss issues. Furthermore, the new yield loss basis can be used to monitor the progress of yield improvement as a discriminate analysis measure for reducing failure patterns (bitmap failures).**

***Index Terms*—Bitmap, cluster analysis, discriminate analysis, principal component analysis (PCA), yield analysis, yield loss space.**

I. Introduction

DURING THE wafer manufacturing process, wafer sorting is the final step to ensure that ICs function well. Only qualified chips are sent on for packaging and further processing. If an IC contains repeated circuit blocks [e.g., embedded static random access memory (SRAM)], bitmap data can be collected as part of chip probe data. Bitmap data collection is a common procedure in SRAM, dynamic random access memory, and Flash memory ICs. Bitmap data records the failing bits of the memory being tested, and it represents the failing counts for different failure patterns recognition. Because specific bitmap failure patterns can be connected to certain semiconductor process failures, several previous studies focus on bitmap failure pattern recognition [2], [8], [9]. The first step of traditional yield analysis approach is to synthesize bitmap data into a Pareto chart. The second step conducts failure analysis to examine selected samples from the Pareto chart to determine failure mechanisms. After modifications are made for the process improvement, the third step generates a new Pareto chart for the second learning cycle. Lastly, the entire procedure is repeated to achieve the yield goal. A few learning cycles and several months of development are usually required to launch new technology (e.g., 45 nm technology).

This paper uses the principal component analysis (PCA) approach to explore the reduced failure space basis, the cluster analysis to generate independent failure events, and discriminate analysis to select objective physical failure analysis (PFA) dice. Using failure analysis, this paper then reveals the IC failure mechanisms and fixes them using process improvement. Results show that the approach in this empirical study demonstrates fewer learning cycles than traditional yield improvement approaches.

SRAM is widely used as a test vehicle in the new generation complementary metal oxide semiconductor (CMOS) process development due to its high front-end transistor density and CMOS-compatible manufacturing process properties. Compared to other kinds of circuits, memory circuits are also easier to use in PFA because failing bits are localized by bitmap testing data. Therefore, SRAM ICs or logic ICs with embedded SRAM are commonly used as monitors in CMOS process development.

Many different mechanisms can induce chip failures. Failure patterns can be categorized as random failures, systematic failures, and repeated failures [5].

This research is organized as follows. Section II introduces the bitmap data and mathematical method, Section III presents examples of bitmap data analysis to demonstrate the procedures of the proposed method, and Section IV presents conclusions and remarks.

II. Bitmap Data and Mathematical Method

This section briefly introduces the bitmap data and the method of generating simulated bitmap data. The second subsection then introduces analytical approaches like PCA, cluster analysis, and discriminate analysis.

A. Bitmap Data

1) *Bitmap Failure Pattern:* Bitmap data records the locations of failing bits in the repeated structures of an IC

Manuscript received November 25, 2008; revised November 12, 2009; accepted June 8, 2010. Date of publication August 26, 2010; date of current version November 3, 2010.

Y.-L. Hsieh is with the Customer Technical Supporting Division, Taiwan Semiconductor Manufacturing Company, Hsinchu 300, Taiwan, and also with the Institute of Management of Technology, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: ylhsieh@tsmc.com).

G.-H. Tzeng, T. R. Lin, and H.-C. Yu are with the Institute of Management of Technology, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: ghtzeng@mail.knu.edu.tw; gtrl@faculty.nctu.edu.tw; chengyu@cc.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSM.2010.2065510

0894-6507/$26.00 © 2010 IEEE

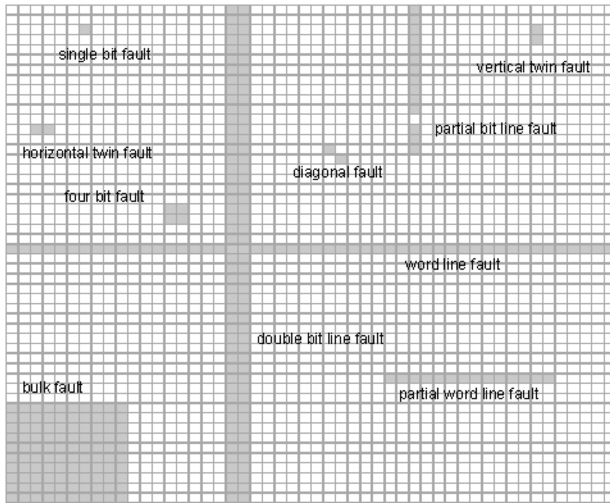

Fig. 1. Bitmap failure pattern recognition.

(e.g., the SRAM). Due to the circuit structure, specific failure mechanisms usually produce particular failure patterns that can be analyzed and categorized. For instance, if four bits share the same via in the circuit layout, they would fail simultaneously by the via open fail. Also, if two bits share the same contact, both would fail if that particular contact fails. Furthermore, failure patterns change with different testing voltages. For example, if the contact fails because of high resistance but no contact open, the results of testing twin bit failure at low voltages and high voltages might be different.

Fig. 1 illustrates typical (SRAM) bitmap failure recognition. Grouping procedures generally begin by processing large-area failure patterns, such as bulk failures, word line failures, and bit line failures. The remaining failing bits are then grouped into four bit failures, twin bit failures, one bit failures, and so on. With further analyses, each of the failure patterns can be classified into several subcategories. For example, wordline failures can be separated into full wordline failures and partial wordline failures, and so on.

2) *Simulated Bitmap Data:* The method in this paper generates bitmap data by the following equation:

$$\mathbf{X}_{n\times p} :\rightarrow \sum_{k=1}^{m} C^k \times \mathbf{D}^k_{n\times 1} \times \mathbf{F}^k_{1\times p} \times f \quad (1)$$

where $\mathbf{X}_{n\times p}$ is a matrix with n row(s) and p column(s), and the element $[x_{vi}]_{n\times p}$ represents the failure bit count of the ith bitmap failure mode ($i = 1, 2, ..., p$) of the vth die ($v = 1, 2, ..., n$), the element $[x_{vi}]_{n\times p}$ can be any nonnegative integer (Ex: if $[x_{12,5}] = 7$, the failure bit count is seven in the 5th failure mode of 12th die). In the simulated example, $n = 500$, $p = 20$, and $m = 6$ (500 dice, 20 failure modes, six yield loss event). The basic idea of (1) is, bitmap data from the kth yield loss event can be constructed by: 1) C^k: fail intensity among different wafers; 2) $\mathbf{D}^k_{n\times 1}$: fail intensity within same wafer but different dice; 3) $\mathbf{F}^k_{1\times p}$: specific bitmap data feature by circuits structure; and 4) f: uncertain factor.

One yield loss event can cause different yield loss results in different wafers (for example, if we have litho machine lens heating issue in a metal layer which makes metal island pattern fail, we can observe same bitmap failures with different "fail intensity," because lens heating is getting worse by process time), so that we use C^k as failure count multipliers to interpret yield loss events' "intensity." The term C^k is a constant of k, which stands for the kth yield loss event, in the model, C^k is assigned as different order for different signal intensities.

TABLE I

C^k VALUES FOR EACH YIELD LOSS EVENT

Event 1 (System Wafer Edge)	Event 2 (System Wafer Edge)	Event 3 (System Wafer Center)	Event 4 (System Localized)	Event 5 (Repeating)	Event 6 (Random)
C^1 = 100	C^2 = 100	C^3 = 100	C^4 = 100	C^5 = 100	C^6 = 1 (low noise scenario) C^6 = 10 (median noise scenario) C^6 = 100 (high noise scenario)

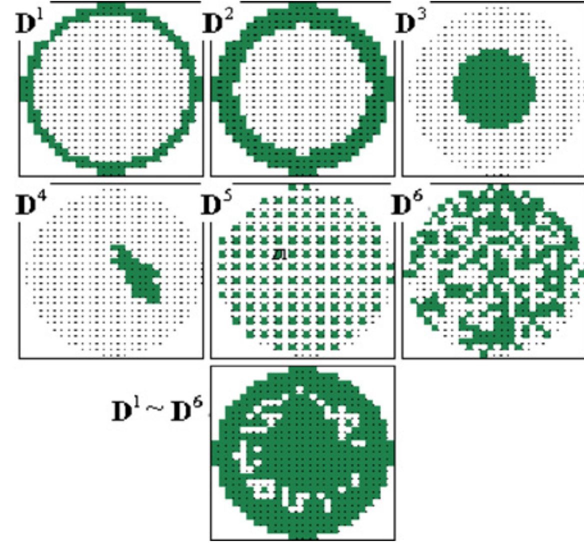

Fig. 2. Wafer map of yield loss events $\mathbf{D}^k_{n\times 1}$.

In Table I, the C^k values function as failure count multipliers for yield loss events. Event-1–Event-5 are signals with three noise scenarios presented by Event 6. When the C^k value of Event 6 (in a low/medium noise scenario) is lower than that of the other events, the signals (Events 1–5) are larger than the noise (Event 6).

$\mathbf{D}^k_{n\times 1}$ is a matrix with n row(s) and one column in Event k, we use "0–1" in $\mathbf{D}^k_{n\times 1}$ to separate those defective/non-defective dices under each yield loss event. For example, $[D_{97}]^4 = 0$ means for 97th die, it is non-defective from yield loss event four. In this paper, we set $m = 6$ (six yield loss events), four systematic yield loss events occur along with one repeated yield loss event and one random yield loss event. Fig. 2 displays the corresponding wafer maps of these six events and combined wafer maps. In the first row of Fig. 2, $\mathbf{D}^1$ and $\mathbf{D}^2$ represent the systematic wafer edge yield losses, while $\mathbf{D}^3$ indicates the systematic wafer center yield loss. In the second row, $\mathbf{D}^4$ represents the systematic localized yield loss, $\mathbf{D}^5$ depicts the repeated yield loss, $\mathbf{D}^6$ shows the random yield loss. Finally, the bottom pattern is a combination of the wafer maps for the six yield loss events above.

A SRAM layout example is shown in Fig. 3, imagining a contact layer process issue inducing a random single contact

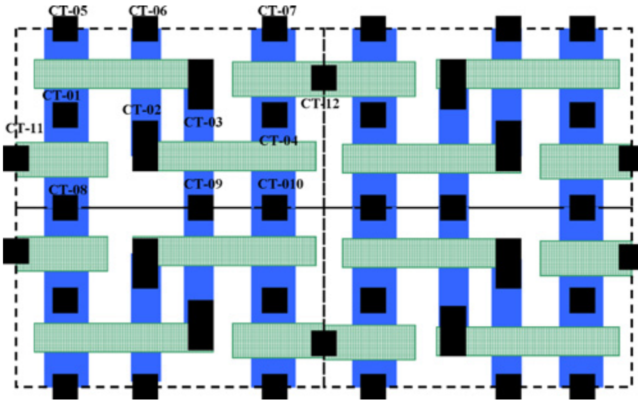

Fig. 3. Four-bit SRAM layout example, OD layer in vertical, POLY layer in horizontal, contact in rectangular [2].

TABLE II
$\mathbf{F}^k_{1\times p}$ VALUES FOR EACH YIELD LOSS EVENT (EXAMPLE)

Fk value list	FM-1	FM-2	FM-3	FM-4	FM-5	FM-6	FM-7	FM-8	FM-9	FM-10	FM-11	FM-12	FM-13	FM-14	FM-15	FM-16	FM-17	FM-18	FM-19	FM-20
Event-1 (systen wafer edge)	0.70	0.10	1.00	1.00	0.15	0.15	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Event-2 (systen wafer edge)	0.20	1.00	0.01	0.01	0.01	0.02	0.20	0.60	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Event-3 (systen wafer center)	0.25	0.25	0.40	0.40	0.45	0.45	0.01	0.01	0.15	0.10	0.50	0.60	1.00	1.00	0.01	0.01	0.01	0.01	0.01	0.01
Event-4 (systen localized)	0.01	0.01	0.01	0.01	0.01	0.01	1.00	1.00	0.80	0.80	0.10	0.10	0.10	0.10	0.20	0.20	0.10	0.10	0.10	0.10
Event-5 (repeating)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	1.00	0.40	0.80	0.30	0.60
Event-6 (random)	random number between [0,1]																			

opening failure, this failure will induce single bit fault if one contact opening among contact-1–contact-4, horizontal twin fault if either contact-11 or contact-12 opening, vertical twin fault if one of contact-5–contact-10 opening. Assuming in a certain yield loss event the contact opening probability is equal to each other, the ratio among single bit fault v.s. horizontal twin bit fault v.s. vertical twin bit fault will be 4:1:3=0.5:0.125:0.375, with a specific signature. Therefore, due to circuit structure characteristics, the ratio of each bitmap failure mode in each event was approximately to a specific signature and be expressed by $\mathbf{F}^k_{1\times p}$, where $\mathbf{F}^k_{1\times p}$ is a matrix with one row and p column(s) in Event k, graded from 0 to 1. TABLE II shows the data used in $\mathbf{F}^k_{1\times p}$.

There are always uncertainties in the testing results. If we test a same wafer in the same tester twice and then compare the results, the data will be close but not 100% the same. The disparities might result from testing marginality, fault marginality, or bitmap fault pattern recognition sensitivity, and so on. In this model, f represents an uncertain factor with a value between 0.8 and 1.2 with conservative estimate.

This paper includes 20 bitmap failure modes and six yield loss events. Table III shows the selected simulated data of **X**. Except for Event 2, the $\mathbf{D}^k_{n\times 1}$values are either 0 or 1. The number 1 signifies that a specific die was influenced by a certain yield loss event; whereas 0 means that the die was not

TABLE III
SELECTED BITMAP DATA LIST
(SEE APPENDIX-D)

Selected Bitmap data	Pass/Fail	Event-1	Event-2	Event-3	Event-4	Event-5	Event-6	FM-1	FM-2	FM-3	FM-4	FM-5	FM-6	FM-7	FM-8	FM-9	FM-10	FM-11	FM-12	FM-13	FM-14	FM-15	FM-16	FM-17	FM-18	FM-19	FM-20
S-1	F	1	0.4	0	0	0	1	87	150	128	121	23	27	37	87	12	14	12	10	12	9	12	13	13	11	10	13
S-2	F	0	0	0	1	1	1	8	10	9	10	12	11	126	123	104	84	20	20	19	19	87	143	58	115	43	80
S-3	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S-4	F	0	0	0	0	0	1	17	78	8	8	11	10	21	49	9	9	11	10	11	8	9	8	8	11	11	8
S-5	F	0	0	0	1	0	0	1	1	1	1	0	0	106	91	78	75	8	11	11	10	22	18	8	10	10	9
S-6	F	0	0	1	0	0	0	29	20	47	41	42	48	1	0	15	11	54	54	100	110	1	1	0	0	0	1
S-7	F	0	0	0	0	0	1	21	70	10	11	11	9	24	37	9	12	9	11	8	8	11	9	9	11	9	10
S-8	F	0	0	1	1	1	0	26	30	36	32	38	42	97	106	111	86	67	59	128	121	68	107	46	79	36	73
S-9	F	0	0	0	0	0	1	19	79	8	11	8	10	20	40	8	10	10	8	8	8	11	8	8	10	8	9
S-10	F	0	0	1	0	0	1	31	35	57	49	63	49	8	11	26	22	66	75	114	119	11	10	10	10	12	10
S-11	F	0	0	1	1	0	0	21	29	48	33	43	40	95	116	92	91	50	76	111	108	18	19	11	9	10	11
S-12	F	0	0.2	0	0	1	0	10	60	0	0	0	1	10	35	0	1	0	0	0	0	43	116	44	93	26	57
S-13	F	0	0	0	0	0	1	22	63	11	10	8	12	20	52	11	8	11	8	11	10	11	11	9	9	9	9
S-14	F	0	0	0	0	0	1	20	65	10	11	11	12	19	48	10	10	8	9	9	8	11	9	9	8	11	11
S-15	P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S-16	F	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	57	102	41	87	26	54
S-17	F	0	0	0	0	0	1	9	9	10	8	10	10	8	9	11	9	9	9	10	11	11	9	8	8	8	9
S-18	F	1	0.6	0	0	0	1	100	202	93	99	27	27	41	107	12	12	9	12	12	9	10	13	11	14	10	11
S-19	F	0	0	0	0	1	1	9	10	9	10	8	8	8	8	9	9	11	9	11	8	55	100	44	103	37	59
S-20	F	0	0	0	0	0	1	8	11	10	11	11	9	9	10	9	8	8	11	8	8	10	10	10	8	10	11

affected. In Event 2, the $\mathbf{D}^k_{n\times 1}$values are set between 0 and 1 to show the intensity of influence that Event 2 has within this wafer. A good die is only produced when all the $\mathbf{D}^k_{n\times 1}$ values are 0, as "P" in the "P/F?" row indicates. The FM1 (bitmap failure mode1) count of Die 1is 110, but 130 for Die 2. Die 1 is affected by Events 1 and Events 2, while Die 2 was affected by Events 1, Events 2, Events 5, and Events 6 (see Appendix D).

B. PCA Approach

PCA is a component of multivariate statistical analysis. PCA was first proposed by Pearson, and then developed by Hotelling. The PCA method is also called the Karhunen-Loève transform or Hotelling transform. PCA is a technique for reducing multidimensional data sets to minimize dimensions of analysis. The dimensions remaining after PCA analysis are mutually independent. In other words, PCA is a linear transformation that converts data to a new coordinate system where the greatest variance of any projection of data lies on the first coordinate (the first principal component), the second greatest variance on the second coordinate, the third greatest variance on the third coordinate, and so forth [1], [4], [6], [7] (see Appendix A).

The following section briefly describes how to use the PCA approach. First, standardize the bitmap raw data to matrix X,

$$x_{vi,s\tan daridized} = \frac{x_{vi,raw} - \bar{x}_i}{\sigma_{x_i}}$$

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{v1} & \cdots & x_{vi} & \cdots & x_{vp} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{ni} & \cdots & x_{np} \end{bmatrix}_{n\times p} = [x_{vi}]_{n\times p}.$$

Second, calculate the correlation matrix $\mathbf{R} = [r_{hj}]_{p\times p}$ where r_{hj} is the correlation coefficient

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1p} \\ r_{21} & 1 & r_{23} & \cdots & r_{2p} \\ r_{31} & r_{32} & 1 & \cdots & r_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & r_{p3} & \cdots & 1 \end{bmatrix}_{p\times p} = [r_{hj}]_{p\times p} \quad (2)$$

of the hth and jth bitmap failure modes, with $p = 20$ in our paper. Third, to calculate Λ, the diagonalization matrix, use

$$\Lambda = \boldsymbol{B}'\boldsymbol{R}\boldsymbol{B} \quad (3)$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}.$$

Λ is a diagonal matrix where the ith diagonal element is the ith eigenvalue of $\mathbf{R} = [r_{hj}]_{p\times p}$ The term $\mathbf{B}$ is a translation matrix between old and new coordinate systems, where the ith column is the ith eigenvector of $\mathbf{R} = [r_{hj}]_{p\times p}$.

In the PCA approach, an important output result called principle factor loading $\mathbf{A} = [a_{ik}]$ can be derived from $\mathbf{B} = [\beta_{ik}]_{p\times p}$ (see Table A2). In other words, the principle factor loading

$$\boldsymbol{A} = [a_{ik}] = [\sqrt{\lambda_k}\beta_{ik}]. \quad (4)$$

The ith eigenvalue λ_k of $[r_{ij}]$ stands for the variance corresponding to kth principle component, so the sum of all eigenvalues is the total variance of $[\boldsymbol{r}_{ij}]$. The variance ratios can be acquired by dividing λ_k by $\sum_{k=1}^{p} \lambda_k = p$ (A7)–(A9).

Another useful result called factor score coefficients may be found by taking m factors that can be derived from dividing $[\beta_{ik}]_{p\times m}$ by $\sqrt{\lambda_k}$.

Factor score coefficients (FSC) are as follows:

$$\text{FSC} = [\beta_{ik} \div \sqrt{\lambda_k}]. \quad (5)$$

Factor scores (FS) can be obtained by multiplying the normalized $\boldsymbol{X}$ by the factor score coefficients

$$\begin{aligned}[\text{FS}]_{n\times m} &= \mathbf{X}_{n\times p} \times \mathbf{B}_{P\times m}/\sqrt{\lambda_k} \\ &= [x_{vi}]_{n\times p} \times [\beta_{ik}/\sqrt{\lambda_k}]_{p\times m} \quad (6)\end{aligned}$$

where variables $i = 1, 2, ..., p$, and samples $v = 1, 2, ..., n$.

C. Cluster and Discriminatory Analysis for Bitmap Failure Patterns

This paper uses clustering to classify a data set into subsets (clusters) so that the data in each subset shares some common features. The cluster analysis employed in this paper uses factor scores and wafer maps.

Due to the nature of the semiconductor process, one die might suffer multiple yield losses. Such dice are not appropriate objectives for failure analysis, because more resources are required to arrive at a failure analysis conclusion. Therefore, this paper proposes using a suitable threshold value of factor score to select PFA objective dice. The objective die should have as more factor scores as possible in specific PC and also have as fewer factor scores as possible for other PCs.

Once the matrix $\mathbf{B}$ is derived from the raw data, (6) can be used to transform the sample's data from bitmap failure space $\mathbf{X}$ to event failure space $\mathbf{Y}$

$$\mathbf{Y} = \mathbf{XB} \quad (7)$$

$$\mathbf{X} = \mathbf{YB}^{-1}. \quad (8)$$

Comparing the spectrum in the event failure space for the original data and transformed data reveals any new yield loss events.

III. Bitmap Data Analysis Example

The conventional yield improvement procedures include: 1) set up the bitmap yield loss Pareto plot; 2) select sample dice and process them with PFA analysis; 3) modify the process recipe to resolve the yield loss events; 4) process new wafers with the new recipe; and 5) continue the above four steps until the yield reaches the desired goal.

Yield loss events are usually mutually independent, which is consistent with the eigenvectors of the bitmap data correlation matrix. Accordingly, Microsoft EXCEL VBA was used to generate the simulated data, PCA analysis, cluster analysis, wafer maps, and so on.

A. Bitmap Data Analysis with Medium Noise

The C^k value of a random yield loss event in a medium noise case is set at 10, which is ten times less than the C^k values of other signals. Table IV lists the top five eigenvalues and factor loadings.

Comparing the factor loading data in Table IV with that in Table II shows that principal component 1 (PC1) is mostly mapped to FM5, FM6 and FM11 to 14 (failure modes 5, 6, 11–14). This reflects the wafer center yield loss event (Event 3 in Table II). PC2 is mainly mapped to FM15 to 20, which reflects the repeated wafer yield loss event (Event 5 in Table II). PC3 is largely mapped to FM1 to 4, FM7 and FM8, which reflects the wafer edge yield loss event (Event 2 in Table II). PC4 is mainly mapped to FM7 to 10, which reflects the system localized yield loss event (Event 4 in Table II). PC5 is mainly mapped to FM2 to 4 and FM8, which reflects the wafer edge yield loss event (Event 1 in Table II).

B. Bitmap Data Analysis with Low Noise

In low noise cases, the C^k value of random yield loss event in Table I is set at 1 for further noise reduction. Repeating the PCA procedures produces the data shown in Table V. These results are very similar to those obtained from bitmap data analysis with medium noise.

C. Bitmap Data Analysis with High Noise

For high noise scenario analysis, the C^k value of random yield loss events in Table I is adjusted to 100 to determine what happens in the PCA analysis when the background

TABLE IV
EIGENVALUES AND CORRESPONDING FACTOR LOADINGS (ROTATED) OF MEDIUM NOISE ANALYSIS

Ck Value of Event 6 = 10	PC1	PC2	PC3	PC4	Community PC1–PC4	PC5
Eigenvalue	6.30	5.87	4.96	2.30		0.24
Proportion	31.5%	29. 3%	24.8%	11 5%		1.2%
Commulative	31.5%	**60.9%**	**85.7%**	97.2%		98.4%
FM1	−0.011	0.020	0.994	−0.022	0.989	−0.025
FM2	−0.185	0.012	0.941	−0.021	0.920	0.266
FM3	0.209	0.026	0.944	−0.027	0.937	−0.228
FM4	0.213	0.018	0.948	−0.032	0.946	−0.204
FM5	0.954	0.027	0.255	0.044	0.978	−0.046
FM6	0.941	0.044	0.301	0.039	0.979	−0.031
FM7	0.080	0.076	0.576	0.795	0.975	0.119
FM8	−0.149	0.028	0.882	0.365	0.934	0.231
FM9	0.532	0.088	−0.043	0.834	0.989	−0.057
FM10	0.455	0.089	0.011	0.879	0.988	−0.048
FM11	0.980	0.037	−0.072	0.132	0.984	0.021
FM12	0.983	0.030	−0.080	0.119	0.987	0.017
FM13	0.982	0.031	−0.100	0.070	0.981	0.021
FM14	0.981	0.017	−0.098	0.079	0.979	0.020
FM15	0.059	0.984	0.010	0.113	0.984	−0.007
FM16	0.014	0.987	0.000	0.006	0.975	−0.001
FM17	0.040	0.989	0.030	0.034	0.981	0.002
FM18	0.000	0.988	0.007	−0.049	0.980	0.006
FM19	0.065	0.978	0.045	0.067	0.967	0.003
FM20	0.021	0.993	0.015	−0.010	0.986	−0.004

TABLE V
EIGENVALUES AND CORRESPONDING FACTOR LOADINGS (ROTATED) OF LOW NOISE ANALYSIS

Ck Value of Event 6 = 1	PC1	PC2	PC3	PC4	Community PC1–PC4	PC5
Eigenvalue	6.31	5.91	4.98	2.30		0.24
Proportion	31.5%	29.6%	24.9%	11.5%		1.2%
Commulative	31.5%	**61.1%**	**86.0%**	97.5%		98.7%
FM1	−0.015	0.011	0.994	−0 032	0.989	0.015
FM2	−0.185	−0.001	0.944	−0.025	0.926	−0.253
FM3	0.206	0.015	0.947	−0.027	0.940	0.222
FM4	0.210	0.021	0.946	−0.034	0.941	0.217
FM5	0.956	0.025	0.257	0.031	0.982	0.045
FM6	0.948	0.027	0.287	0.036	0.984	0.032
FM7	0.077	0.069	0.586	0.786	0.973	−0.114
FM8	−0.149	0.035	0.885	0.355	0.933	−0.231
FM9	0.530	0.085	−0.044	0.834	0.986	0.052
FM10	0.445	0.069	0.005	0.887	0.990	0.043
FM11	0.979	0.032	−0.078	0.128	0 982	−0.021
FM12	0.979	0.033	−0.084	0.137	0.986	−0.017
FM13	0.985	0.026	−0.099	0.080	0.987	−0.020
FM14	0.985	0.017	−0.100	0.073	0.986	−0.020
FM15	0.048	0.986	0.008	0.099	0.984	0.009
FM16	0.011	0.992	−0.001	−0.004	0.985	0.001
FM17	0.034	0.991	0.028	0.022	0.985	−0.013
FM18	0.008	0.993	−0.001	−0.035	0.987	−0.006
FM19	0.051	0.989	0.038	0.067	0.987	0.005
FM20	0.020	0.993	0.010	−0.011	0.987	0.004

TABLE VI
EIGENVALUES AND CORRESPONDING FACTOR LOADINGS (ROTATED) OF HIGH NOISE ANALYSIS

Ck Value of Event 6 = 100	PC1	PC2	PC3	Community PC1–PC4	PC4	PC5
Eigenvalue	14.69	2.77	1.36		0.41	0.32
Proportion	73.5%	13.8%	6.8%		2.1%	1.6%
Commulative	73.5%	**87.3%**	**94.1%**		96.2%	97.8%
FM1	0.683	0.710	0.016	0.972	0.100	0.050
FM2	0.391	0.898	−0.068	0.964	−0.009	−0.025
FM3	0.766	0.582	0.088	0.934	0.171	0.002
FM4	0.754	0.604	0.091	0.941	0.174	−0.005
FM5	0.969	0.069	0.142	0.963	0.099	−0.037
FM6	0.970	0.090	0.132	0.965	0.094	−0.023
FM7	0.868	0.333	0.050	0.867	−0.328	0.091
FM8	0.611	0.732	−0.023	0.910	−0.278	0.018
FM9	0.947	−0 016	0.126	0.913	−0.249	0.057
FM10	0.950	0.014	0.109	0.914	−0.236	0.108
FM11	0.970	−0 064	0.150	0.967	0.049	−0.059
FM12	0.961	−0.094	0.169	0.961	0.072	−0.107
FM13	0.883	−0.177	0.241	0.869	0.085	−0.332
FM14	0.881	−0.193	0.253	0.877	0.080	−0.320
FM15	0.919	0.037	−0.343	0.963	−0.014	0.104
FM16	0.777	0.010	−0.613	0.980	−0.003	−0.048
FM17	0.934	0.061	−0.293	0.962	0.061	0.133
FM18	0.836	0.018	−0.530	0.980	0.025	0.018
FM19	0.946	0.061	−0.216	0.944	0.008	0.167
FM20	0.884	0.040	−0.435	0.972	0.037	0.081

noise increases. When the noise intensity matches the signal intensity, the failure counts of systematic failure events roughly equal the number of random failure events. Yield loss events cannot be completely decoupled by PCA, as Table VI shows. This result shows that PC1 is a random failure event, and only three out of five events can be analyzed: wafer center, localized, and repeated failure events (PC2–PC4). However, in real situations, the failure counts of random yield loss events are lower than systematic and repeated yield loss events. As a result, high noise cases rarely occur.

D. *Cluster Analysis*

The details of yield loss events are usually unavailable for both the event (number) count and event bitmap failure mode distribution. Therefore, factor score data is required for deciding objective sample dice of the PFA. The factor scores can be easily obtained, as (6) indicates.

Fig. 4 plots the PC1 factor scores and corresponding $\mathbf{D}_{n\times 1}^{k}$ values at a medium noise level for 500 simulated samples. According to the plot, the factor score data varies significantly among various $\mathbf{D}_{n\times 1}^{k}$ values (0 or 1). As Fig. 5 shows, the factor score map of PC1 is very similar to the $\mathbf{D}_{n\times 1}^{k}$ map of Event 3. The threshold values of the gray-level parts of the factor score map are 1.88, 1.28, and 0.84, respectively. These values correspond to roughly 97%, 90%, 80% of the cumulative probability.

Fig. 6 shows that the factor score map of PC2 is similar to the $\mathbf{D}_{n\times 1}^{k}$ map of Event 5. The original yield loss events repeat the defects possibly caused by the mask's faults. Except for the wafer center area, the repeated defects of yield loss events can be duplicated in PC2's factor score map. This is because

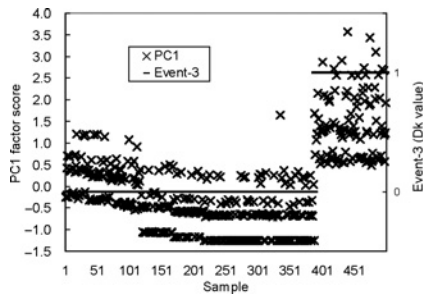

Fig. 4. Factor score plot for each sample with corresponding $\mathbf{D}^k_{n\times 1}$ value (medium noise, PC1).

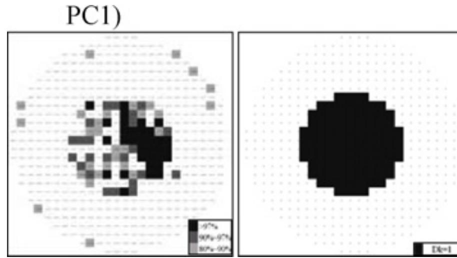

Fig. 5. Factor score map for PC1 (left) and $\mathbf{D}^k_{n\times 1}$ map (right) for Event 3 (medium noise).

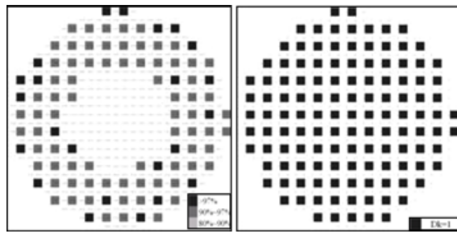

Fig. 6. Factor score map for PC2 (left) and $\mathbf{D}^k_{n\times 1}$ map (right) for Event 5 (medium noise).

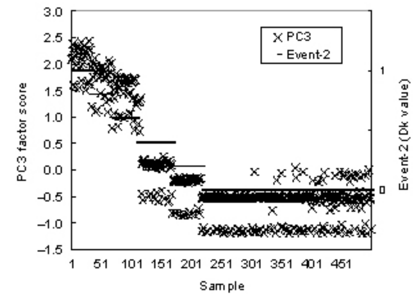

Fig. 7. Factor score plot for each sample with corresponding $\mathbf{D}^k_{n\times 1}$value (medium noise, PC3).

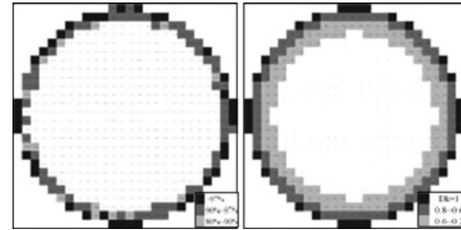

Fig. 8. Factor score map for PC3 (left) and $\mathbf{D}^k_{n\times 1}$ map (right) for Event 2 (medium noise).

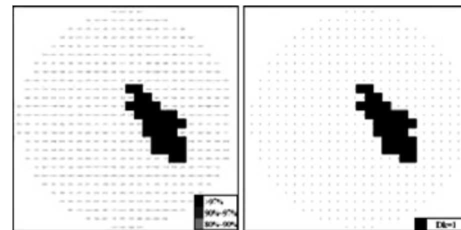

Fig. 9. Factor score map for PC4 (left) and $\mathbf{D}^k_{n\times 1}$ map (right) for Event 4 (medium noise).

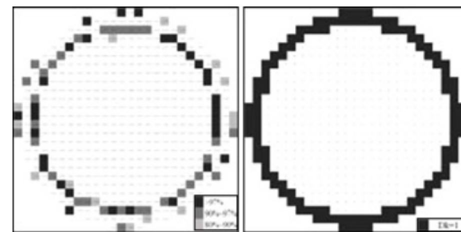

Fig. 10. Factor score map for PC5 (left) and $D^k_{n\times 1}$ map (right) for Event 1 (medium noise).

the wafer center signal might be biased from the wafer center yield loss event (Event 3 in PCA analysis).

In Event 2, the $\mathbf{D}^k_{n\times 1}$ values are set at 0, 0.2, 0.4, 0.6, 0.8, and 1.0 to represent varying degrees of yield loss. According to Fig. 7, the magnitude of PC3 is correlated to the $\mathbf{D}^k_{n\times 1}$ values of Event 2, which supports the knowledge that greater factor score values correspond to more serious degrees of yield loss.

Fig. 8 shows that PC3 represents Event 2, Fig. 9 shows that PC4 represents Event 4, and Fig. 10 indicates that PC5 stands for Event 1. The eigenvalue of PC5 is much less than that of the other PCs, so the connection between PC5 and Event 1 is not obvious.

Based on this discussion above, the objective die ν should be on the right-hand side of the factor score probability distribution for the given factor scores of specific PC. Simultaneously, the objective die should also have as fewer factor scores as possible for other PCs.

Equation set (7) shows the cluster analysis proposed in this paper. Appendix B is the factor score summary of medium noise scenario (C^k = 10), with gray background mark for $P\{[FS]_k \leq \text{threshold}_k\} = 97\%$ and an underlined mark for $P\{[FS]_k \leq \text{threshold}^*_k\} = 70\%$. For example, die-482 is a proposed candidate of PC1, die-274 is a proposed candidate of PC2, die-008 is a proposed candidate of PC3, and die-387 is a proposed candidate of PC4. The criteria 97% and 70% are not invariant, they depend on how many dice can be supported by PFA resource

$$\begin{cases} [FS]_{\nu k} \geq \text{threshold}_k \quad \text{for } k = \eta \\ [FS]_{\nu k} \leq \text{threshold}^*_k \quad \text{for } k \neq \eta. \\ \quad k=1,2,.\eta..,6 \quad \nu=1,2,\ldots,500 \end{cases} \tag{9}$$

E. Empirical Study

An empirical study in the 65 nm SoC device yield improvement trend chart is shown in Fig. 11. In this example, ten

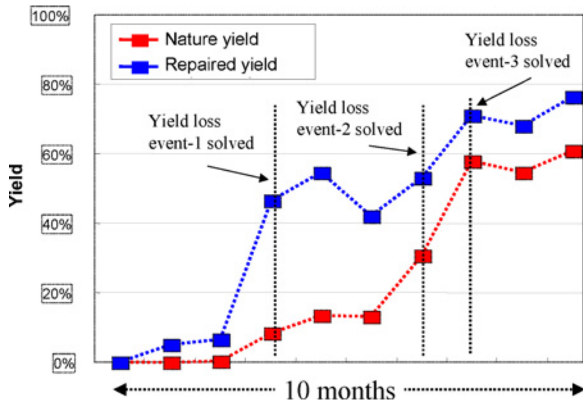

Fig. 11. Selected 65 nm device yield improvement trend chart.

TABLE VII
EIGENVALUES AND CORRESPONDING FACTOR LOADINGS IN THE EMPIRICAL STUDY

	Factor-1	Factor-2	Factor-3	Factor-4	Factor-5	Factor-6
eigenvalues	5.1	1.6	1.3	1.2	0.8	0.6
bit map data	Factor Loadings					
FM-01	0.78	-0.20	-0.28	0.43	-0.07	-0.08
FM-02	0.35	-0.26	-0.31	0.75	0.14	0.27
FM-03	0.74	-0.30	0.11	-0.01	0.13	-0.47
FM-04	0.55	0.12	0.65	0.25	-0.13	-0.18
FM-05	0.57	0.76	-0.21	0.02	0.04	-0.07
FM-06	0.77	-0.13	0.18	-0.08	0.49	-0.09
FM-07	0.77	-0.04	-0.25	-0.14	-0.43	0.00
FM-08	0.82	-0.11	-0.12	-0.14	-0.23	0.03
FM-09	0.68	-0.36	-0.06	-0.41	-0.24	0.18
FM-10	0.40	0.19	0.72	0.21	-0.20	0.32
FM-11	0.53	0.79	-0.22	-0.03	0.08	0.03
FM-12	0.68	-0.10	0.07	-0.36	0.44	0.35
defective rate	4%	5%	1%	12%	1%	6%

months were taken to improve the yield from 0% to 80% by traditional approach.

The traditional procedure is: 1) build up a yield loss Pareto from tested data; 2) 80–20 rule, choose top failure dice for PFA, usually only couple of dice are chosen due to limited PFA resource; 3) draw up a process improvement strategy according to PFA results; and 4) after new wafer processed by improved process condition, repeat procedures (1)–(3).

If there were two yield loss events (named event A and event B), event A causes hundreds of failure bits per die, and event B only causes couple of failure bits per die. Most likely, we could not dig out event B in the first learning iteration by traditional procedure.

Table VII is the PCA results of the 0% yield wafer of mentioned 65 nm device. In this paper, we add defective rate from factor score to improve PCA in practice (in this case, factor score >0.7 is treated as defective dice). Compared to the PCA data and empirical results, PCA results factor-2, factor-4, factor-6 can match yield loss Event 2, Event 3, Event 1, respectively. PCA results factor-1 can be treated as random defect, because factor-1 can cause whole the bit map failure modes FM01–FM12. We did not put resource on Factor-3/5, because the defective rate is as low as 1%.

F. Discussion and Implementation

From the simulated analysis of medium and low noise cases, using the PCA based approach discusses all the principal components, and four out of the five eigenvalues are greater than one. From the empirical example, which introduces defective rate to enhance the PCA results in practice, one can shorten the yield learning iterations. Based on this paper, objective dice of the PFA for yield improvement can be selected after PCA is implemented using the following proposed procedures:

1) 80–20 rule, calculating the defect rate of each factor based on factor scores which are greater than a suitable value (ex: 0.7);
2) from high defective PCs, choosing the objective dice as higher factor score as possible with corresponding PC and also as lower factor score as possible in other PCs;
3) drawing up a process improvement strategy according to PFA results on objective dice;
4) after new wafer processed by improved process condition, repeat procedures (1)–(3).

Based on the analysis of the three different noise scenarios, the PCA based approach is influenced when the noise intensity equals the signal intensity (i.e., the high noise case). However, most yield loss events can be detected when the noise intensity is smaller than the signal intensity.

Wafer map overlaps between yield loss events indeed affecting the results produced by PCA. For example, Event 3 and Event 5 have similar wafer center area map, and this overlap influences the factor score maps of PC1 and PC2.

It is not practicable to have PFA in whole defective dice, usually several dice are selected as PFA objectives for one yield learning iteration. Usually, several learning iterations are needed for yield improvement and it really consumes both time (one iteration—three months) and cost (need many wafers). The proposed "PCA+ defective rate" analysis is a practical methodology for shortened yield learning iteration, compared to traditional Pareto rules.

IV. CONCLUSION

The process of new technology yield improvement can take up to one year to complete [3]. The bottlenecks in this process are: 1) not all yield loss events can be uncovered in the first analysis; 2) after the modifications, two or three months of process time is required to verify the yield results; and 3) the PFA method is a resource-limited and time-consuming approach.

However, the bitmap data analysis method proposed in this paper uses the "PCA based + defective rate" approach to greatly reduce the yield learning cycle time without requiring additional resources. Only a desk-top computer, the related software, and a little time is required to conduct this analysis.

Although previous studies present numerous data mining approaches [1]–[6], [8], [11], none of them can decouple the yield loss events considering wafer maps and signal intensity. Only the PCA based approach can decouple these kinds of failures.

Once the bitmap data is analyzed, the principal components can be used as the basis of failure process space. The following techniques are suggested for semiconductor manufacturing yield management: 1) the goal of yield improvement should objective not only systematic or repeated failure events, but also random failure events; 2) since the failure process space of a wafer analyzed by PCA can be established, the problems

of another wafer with similar failure modes can be disclosed in minutes without any traditional analysis; and 3) the basis of failure process space can be extended by adding new failure modes. Eventually, a complete version of failure process space for a specific technology node can be built (ex: 0.13 μm CMOS logic low power process) to improve manufacturing knowledge management or technology transfer.

APPENDIX A
CONCEPTS OF PRINCIPLE COMPONENT

TABLE A.1
BITMAP DATA SET MATRIX AND SAMPLE SCORES FOR PCA

Samples	Variables					Sample scores				
	x_1	$\cdots$	x_i	$\cdots$	x_p	y_1		y_k		y_m
1	x_{11}	$\cdots$	x_{1i}	$\cdots$	x_{1p}	y_{11}	$\cdots$	y_{1i}	$\cdots$	y_{1m}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$		$\vdots$
v	x_{v1}	$\cdots$	x_{vi}	$\cdots$	x_{vp}	y_{v1}	$\cdots$	y_{vi}	$\cdots$	y_{vm}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$		$\vdots$		$\vdots$
n	x_{n1}	$\cdots$	x_{ni}	$\cdots$	x_{np}	y_{n1}	$\cdots$	y_{ni}	$\cdots$	y_{nm}
$\bar{x}$	$\bar{x}_1$	$\cdots$	$\bar{x}_i$	$\cdots$	$\bar{x}_p$	–	$\cdots$	–	$\cdots$	–
s	s_1	$\cdots$	s_i	$\cdots$	s_p	–	$\cdots$	–	$\cdots$	–

TABLE A.2
PRINCIPLE FACTOR LOADING

Variables	Principle y_1		Principle y_k		Principle y_m	Community
x_1	$\sqrt{\lambda_1}\beta_{11}$	$\cdots$	$\sqrt{\lambda_k}\beta_{1k}$	$\cdots$	$\sqrt{\lambda_m}\beta_{1m}$	h_1
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_i	$\sqrt{\lambda_1}\beta_{i1}$	$\cdots$	$\sqrt{\lambda_k}\beta_{ik}$	$\cdots$	$\sqrt{\lambda_m}\beta_{im}$	h_i
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_p	$\sqrt{\lambda_1}\beta_{p1}$	$\cdots$	$\sqrt{\lambda_k}\beta_{pk}$	$\cdots$	$\sqrt{\lambda_m}\beta_{pm}$	h_p
Eigenvalue	λ_1	$\cdots$	λ_k	$\cdots$	λ_m	–
Contri. rate	λ_1/p	$\cdots$	λ_k/p	$\cdots$	λ_m/p	.
Accu.con. rate	λ_1/p	$\cdots$	$\sum_{k=1}^{k}\lambda_k/p$	$\cdots$	$\sum_{k=1}^{m}\lambda_m/p$	

where $a_{ik} = \sqrt{\lambda_k}\beta_{ik}$,i.e., $A = [a_{ik}] = [\sqrt{\lambda_k}\beta_{ik}]$, $h_i = \sum_{k=1}^{m} a_{ik}^2$ and $\sqrt{\lambda_k}\beta_{ik}$ and $\lambda_k = \sum_{i=1}^{p} a_{ik}^2$.

A. *Basic Concepts of Principle Component*

$$Y = \beta_1 x_1 + \cdots + \beta_i x_i \cdots + \beta_p x_p = \boldsymbol{\beta}' \boldsymbol{x}. \tag{A1}$$

Let the synthetic index $Y = \boldsymbol{\beta}'\boldsymbol{x}$ of standardized bitmap variables vector $\boldsymbol{x} = (x_1, \cdots, x_i, \cdots, x_p)$ be the greatest variance, i.e., maximize $Var(Y) = Var(\boldsymbol{\beta}'\boldsymbol{x}) = \boldsymbol{\beta}'\boldsymbol{R}\boldsymbol{\beta}$. Let the standardized bitmap data set matrix $\boldsymbol{X}$ be

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{v1} & \cdots & x_{vi} & \cdots & x_{vp} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{ni} & \cdots & x_{np} \end{bmatrix}_{n\times p} = [x_{vi}]_{n\times p}. \tag{A2}$$

APPENDIX-B: SELECTIVE FACTOR SCORES (C^k = 10)

TABLE A.3
FACTOR SCORES (C^k=10) OF DIE-1 DIE-250

DIE	PC1	PC2	PC3	PC4	DIE	PC1	PC2	PC3	PC4
DIE-001	-0.25	-0.12	2.09	0.03	DIE-126	-0.46	-0.13	0.13	0.51
DIE-002	-0.24	-0.13	2.12	-0.07	DIE-127	-0.48	-0.13	0.11	0.55
DIE-003	0.42	0.08	2.38	-0.04	DIE-128	-0.50	-0.17	0.06	0.50
DIE-004	0.71	1.94	1.61	-0.46	DIE-129	-0.50	-0.15	0.13	0.51
DIE-005	-0.25	-0.13	2.16	-0.31	DIE-130	0.38	1.83	-0.50	0.31
DIE-006	0.38	0.07	2.34	-0.19	DIE-131	-1.06	-0.30	0.10	0.46
DIE-007	0.67	1.82	1.56	-0.35	DIE-132	-0.20	1.64	-0.45	0.21
DIE-008	-0.20	-0.15	2.17	-0.18	DIE-133	-1.07	-0.31	0.06	0.48
DIE-009	0.35	0.05	1.98	0.12	DIE-134	-1.07	-0.31	0.05	0.51
DIE-010	-0.18	-0.12	2.33	-0.11	DIE-135	-1.05	-0.29	0.13	0.55
DIE-011	-0.18	-0.11	2.22	0.08	DIE-136	-1.06	-0.29	0.14	0.43
DIE-012	0.40	0.13	2.31	0.06	DIE-137	-0.47	-0.12	0.07	0.56
DIE-013	-0.14	-0.13	2.35	-0.16	DIE-138	-0.15	1.74	-0.51	0.26
DIE-014	0.70	1.88	1.55	-0.25	DIE-139	-1.05	-0.30	0.07	0.51
DIE-015	0.32	0.01	2.21	-0.22	DIE-140	-0.14	1.79	-0.56	0.13
DIE-016	0.73	1.80	1.95	-0.49	DIE-141	-1.07	-0.29	0.10	0.42
DIE-017	0.47	0.07	2.42	0.11	DIE-142	-1.07	-0.29	0.13	0.45
DIE-018	1.21	1.99	1.69	-0.30	DIE-143	-0.20	1.65	-0.43	0.19
DIE-019	0.34	0.02	2.07	0.20	DIE-144	-0.50	-0.13	0.11	0.56
DIE-020	-0.29	-0.13	2.00	0.11	DIE-145	-0.48	-0.11	0.16	0.68
DIE-021	-0.22	-0.11	2.21	-0.10	DIE-146	0.37	1.78	-0.58	0.23
DIE-022	0.41	0.11	2.34	0.26	DIE-147	-0.50	-0.11	0.10	0.51
DIE-023	-0.22	-0.16	2.15	-0.11	DIE-148	-0.47	-0.14	0.11	0.56
DIE-024	0.37	0.04	2.26	0.07	DIE-149	-1.08	-0.31	0.04	0.46
DIE-025	0.68	1.79	1.50	-0.34	DIE-150	-1.06	-0.28	0.13	0.52
DIE-026	0.38	0.10	2.19	-0.02	DIE-151	-0.25	1.51	-0.40	0.28
DIE-027	-0.19	-0.15	2.21	-0.23	DIE-152	-1.07	-0.28	0.13	0.42
DIE-028	1.18	1.94	1.62	-0.21	DIE-153	-1.08	-0.31	0.03	0.39
DIE-029	-0.18	-0.10	2.28	0.22	DIE-154	-0.46	-0.15	0.25	0.58
DIE-030	1.16	1.82	1.67	-0.27	DIE-155	-0.46	-0.13	0.17	0.54
DIE-031	-0.12	-0.10	2.43	0.14	DIE-156	0.46	1.99	-0.54	0.22
DIE-032	0.38	0.04	2.27	-0.23	DIE-157	-1.07	-0.30	0.06	0.47
DIE-033	1.20	1.90	1.60	-0.44	DIE-158	-1.06	-0.29	0.13	0.44
DIE-034	0.28	-0.03	1.93	-0.12	DIE-159	-0.51	-0.16	0.09	0.51
DIE-035	0.51	1.74	1.20	-0.26	DIE-160	-0.49	-0.14	0.15	0.54
DIE-036	0.29	0.03	2.07	-0.16	DIE-161	-0.45	-0.07	0.16	0.52
DIE-037	0.38	0.05	2.00	-0.02	DIE-162	0.37	1.85	-0.59	0.11
DIE-038	-0.32	-0.21	1.74	-0.21	DIE-163	0.50	2.05	-0.59	0.24
DIE-039	-0.29	-0.18	1.96	-0.32	DIE-164	-0.48	-0.07	0.18	0.52
DIE-040	0.30	0.03	2.11	-0.12	DIE-165	-1.07	-0.30	0.05	0.41
DIE-041	0.60	1.88	1.10	-0.56	DIE-166	-1.05	-0.28	0.13	0.48
DIE-042	-0.33	-0.16	1.88	-0.22	DIE-167	-1.07	-0.29	0.09	0.43
DIE-043	0.21	-0.02	1.83	-0.10	DIE-168	0.46	1.93	-0.46	0.35
DIE-044	0.24	0.01	1.92	-0.29	DIE-169	-0.31	1.59	-0.82	0.04
DIE-045	1.20	2.12	1.13	-0.57	DIE-170	-1.18	-0.36	-0.20	0.33
DIE-046	0.28	0.01	1.91	-0.39	DIE-171	-0.57	-0.15	-0.21	0.44
DIE-047	0.23	-0.03	1.74	-0.06	DIE-172	-0.59	-0.19	-0.17	0.42
DIE-048	-0.31	-0.19	1.90	-0.30	DIE-173	-1.18	-0.37	-0.21	0.28
DIE-049	-0.34	-0.20	1.79	-0.37	DIE-174	0.29	1.83	-0.81	0.14
DIE-050	1.19	1.94	1.45	-0.51	DIE-175	-0.27	1.68	-0.81	0.03
DIE-051	0.30	-0.04	1.95	-0.14	DIE-176	-0.61	-0.18	-0.20	0.38
DIE-052	0.21	-0.01	1.80	-0.14	DIE-177	-1.17	-0.36	-0.19	0.31
DIE-053	0.34	0.02	1.95	-0.18	DIE-178	-0.58	-0.19	-0.20	0.33
DIE-054	0.28	-0.07	1.88	-0.19	DIE-179	-1.18	-0.37	-0.23	0.29
DIE-055	-0.32	-0.22	1.73	-0.16	DIE-180	-0.59	-0.21	-0.13	0.45
DIE-056	0.61	1.86	1.40	-0.51	DIE-181	0.33	1.84	-0.84	0.03
DIE-057	-0.34	-0.19	1.86	-0.27	DIE-182	-0.60	-0.18	-0.16	0.37
DIE-058	0.22	-0.06	1.74	-0.22	DIE-183	-0.62	-0.15	-0.18	0.40
DIE-059	0.56	1.83	1.29	-0.63	DIE-184	-0.62	-0.21	-0.20	0.40
DIE-060	0.24	-0.03	1.80	-0.04	DIE-185	-0.58	-0.19	-0.19	0.42
DIE-061	-0.29	-0.22	1.81	-0.36	DIE-186	-0.30	1.64	-0.86	-0.03
DIE-062	1.15	2.02	1.09	-0.44	DIE-187	-1.17	-0.36	-0.18	0.33
DIE-063	-0.26	-0.19	2.07	-0.41	DIE-188	-1.18	-0.37	-0.22	0.33
DIE-064	-0.28	-0.18	2.00	-0.10	DIE-189	-0.28	1.66	-0.84	0.02
DIE-065	0.60	1.91	1.32	-0.39	DIE-190	-1.17	-0.36	-0.19	0.34
DIE-066	0.31	-0.06	1.83	-0.17	DIE-191	-0.57	-0.17	-0.22	0.41
DIE-067	-0.35	-0.21	1.66	-0.29	DIE-192	-0.61	-0.22	-0.19	0.44
DIE-068	-0.25	-0.19	2.01	-0.24	DIE-193	-1.18	-0.36	-0.21	0.32
DIE-069	0.34	1.61	0.78	-0.49	DIE-194	-0.28	1.66	-0.84	0.00
DIE-070	0.17	-0.11	1.52	-0.34	DIE-195	-0.60	-0.19	-0.23	0.36
DIE-071	0.52	1.80	1.07	-0.77	DIE-196	-0.59	-0.21	-0.17	0.29
DIE-072	0.13	-0.13	1.35	-0.25	DIE-197	0.26	1.69	-0.80	0.08
DIE-073	0.15	-0.04	1.53	-0.42	DIE-198	-0.58	-0.17	-0.15	0.40
DIE-074	0.18	-0.06	1.62	-0.29	DIE-199	-0.58	-0.20	-0.18	0.36
DIE-075	0.16	-0.05	1.47	-0.13	DIE-200	-0.62	-0.21	-0.17	0.33
DIE-076	0.20	-0.07	1.74	-0.40	DIE-201	-0.30	1.61	-0.83	0.01
DIE-077	0.25	-0.02	1.82	-0.39	DIE-202	-1.17	-0.36	-0.18	0.31
DIE-078	0.43	1.76	0.81	-0.72	DIE-203	-0.60	-0.19	-0.19	0.40
DIE-079	-0.37	-0.24	1.70	-0.46	DIE-204	-0.62	-0.18	-0.16	0.33
DIE-080	0.63	1.91	1.01	-0.76	DIE-205	-0.61	-0.20	-0.17	0.36
DIE-081	-0.40	-0.27	1.57	-0.44	DIE-206	-0.57	-0.20	-0.18	0.41
DIE-082	-0.41	-0.21	1.76	-0.56	DIE-207	0.35	1.91	-0.85	0.08
DIE-083	0.19	-0.04	1.70	-0.39	DIE-208	-1.17	-0.36	-0.17	0.33
DIE-084	-0.43	-0.20	1.63	-0.28	DIE-209	-0.61	-0.18	-0.15	0.38
DIE-085	0.40	1.54	1.00	-0.81	DIE-210	-1.18	-0.37	-0.22	0.27
DIE-086	0.19	-0.06	1.62	-0.40	DIE-211	-0.63	-0.18	-0.23	0.32
DIE-087	0.37	1.51	0.98	-0.67	DIE-212	0.21	1.63	-0.73	0.11
DIE-088	0.12	-0.12	1.43	-0.25	DIE-213	-0.58	-0.18	-0.18	0.40
DIE-089	0.23	-0.07	1.72	-0.35	DIE-214	-0.31	1.59	-0.81	0.00
DIE-090	0.19	-0.06	1.72	-0.28	DIE-215	-0.60	-0.18	-0.17	0.42
DIE-091	-0.44	-0.23	1.64	-0.43	DIE-216	-0.57	-0.17	-0.16	0.42
DIE-092	-0.39	-0.24	1.73	-0.38	DIE-217	-1.25	-0.42	-0.53	0.12
DIE-093	-0.37	-0.25	1.69	-0.42	DIE-218	-0.36	1.60	-1.15	-0.15
DIE-094	-0.40	-0.26	1.65	-0.52	DIE-219	-0.70	-0.27	-0.49	0.15
DIE-095	0.17	-0.05	1.72	-0.40	DIE-220	-1.25	-0.42	-0.53	0.12
DIE-096	-0.45	-0.27	1.43	-0.28	DIE-221	0.16	1.65	-1.09	-0.08
DIE-097	-0.48	-0.25	1.45	-0.40	DIE-222	-0.68	-0.27	-0.51	0.21
DIE-098	0.58	1.91	1.06	-0.73	DIE-223	-0.66	-0.24	-0.50	0.20
DIE-099	-0.46	-0.27	1.46	-0.24	DIE-224	-0.67	-0.23	-0.50	0.27
DIE-100	1.08	1.98	1.05	-0.44	DIE-225	0.26	1.82	-1.16	-0.10
DIE-101	-0.37	-0.25	1.66	-0.43	DIE-226	-1.25	-0.42	-0.53	0.12
DIE-102	-0.41	-0.25	1.68	-0.49	DIE-227	-0.33	1.65	-1.17	-0.16
DIE-103	-0.46	-0.26	1.35	-0.36	DIE-228	-1.25	-0.42	-0.53	0.12
DIE-104	0.50	1.87	0.91	-0.71	DIE-229	-0.67	-0.21	-0.53	0.19
DIE-105	-0.43	-0.23	1.73	-0.41	DIE-230	-0.64	-0.23	-0.51	0.18
DIE-106	0.19	-0.06	1.69	-0.33	DIE-231	-0.67	-0.28	-0.49	0.20
DIE-107	0.58	1.95	0.76	-0.92	DIE-232	-1.25	-0.42	-0.53	0.12
DIE-108	0.20	-0.03	1.70	-0.26	DIE-233	-0.67	-0.22	-0.50	0.22
DIE-109	-0.54	-0.33	1.22	-0.63	DIE-234	-1.25	-0.42	-0.53	0.12
DIE-110	0.10	-0.12	1.29	-0.51	DIE-235	-1.25	-0.42	-0.53	0.12
DIE-111	0.04	-0.12	1.28	-0.65	DIE-236	-1.25	-0.42	-0.53	0.12
DIE-112	0.91	1.63	0.88	-0.98	DIE-237	-0.68	-0.24	-0.50	0.19
DIE-113	-0.52	-0.31	1.28	-0.58	DIE-238	0.25	1.84	-1.14	-0.07
DIE-114	0.03	-0.13	1.31	-0.61	DIE-239	-1.25	-0.42	-0.53	0.12
DIE-115	0.31	1.49	0.73	-0.96	DIE-240	-0.66	-0.23	-0.49	0.21
DIE-116	0.11	-0.15	1.36	-0.57	DIE-241	-0.65	-0.26	-0.50	0.24
DIE-117	0.39	1.80	-0.45	0.30	DIE-242	0.21	1.70	-1.11	-0.03
DIE-118	-0.46	-0.10	0.12	0.54	DIE-243	-0.41	1.50	-1.13	-0.16
DIE-119	-0.20	1.67	-0.57	0.13	DIE-244	-0.69	-0.24	-0.49	0.23
DIE-120	-0.45	-0.12	0.12	0.52	DIE-245	-1.25	-0.42	-0.53	0.12
DIE-121	-1.06	-0.28	0.16	0.41	DIE-246	-0.68	-0.23	-0.49	0.28
DIE-122	-0.50	-0.08	0.19	0.52	DIE-247	-0.66	-0.26	-0.50	0.17
DIE-123	-0.48	-0.14	0.18	0.51	DIE-248	-1.25	-0.42	-0.53	0.12
DIE-124	-1.08	-0.28	0.14	0.49	DIE-249	-0.67	-0.26	-0.49	0.21
DIE-125	-1.07	-0.30	0.08	0.45	DIE-250	-0.36	1.60	-1.15	-0.18

The correlation matrix $\boldsymbol{R} = [r_{ij}]_{p \times p}$ is a variance-covariance matrix by the standardized bitmap data set matrix $\boldsymbol{X} = [x_{iv}]_{n \times p}$, where $\boldsymbol{R} = [r_{ij}]_{p \times p}$ is the correlation coefficient matrix

$$\boldsymbol{R} = [r_{ij}]_{p \times p} = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1n} \\ r_{21} & 1 & r_{23} & \cdots & r_{2n} \\ r_{31} & r_{32} & 1 & \cdots & r_{3n} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ r_{n1} & r_{n2} & r_{n3} & \cdots & 1 \end{bmatrix}. \tag{A3}$$

Find the eigenvalue and respective eigenvalue to eigenvector as follows:

$$\text{Maximize } Var(Y) = \boldsymbol{\beta}' \boldsymbol{R} \boldsymbol{\beta} \tag{A4}$$

subject to

$$\sum_{i=1}^{p} \beta_i^2 = \boldsymbol{\beta}' \boldsymbol{\beta} = 1.$$

Take the Lagrange under equality constraint $\boldsymbol{\beta}' \boldsymbol{\beta} = 1$ for maximizing objective function $Var(Y) = \boldsymbol{\beta}' \boldsymbol{R} \boldsymbol{\beta}$, and then see that

$$Q = \boldsymbol{\beta}' \boldsymbol{R} \boldsymbol{\beta} - \lambda(\boldsymbol{\beta}' \boldsymbol{\beta} - 1) \rightarrow \text{ Maximize} \tag{A5}$$

$$\begin{aligned} \frac{\partial Q}{\partial \boldsymbol{\beta}} &= \frac{\partial}{\partial \boldsymbol{\beta}} [\boldsymbol{\beta}' R \boldsymbol{\beta} - \lambda(\boldsymbol{\beta}' \boldsymbol{\beta} - 1)] \\ &= 2R\boldsymbol{\beta} - 2\lambda\boldsymbol{\beta} = 0 \Rightarrow R\boldsymbol{\beta} = \lambda\boldsymbol{\beta}. \end{aligned} \tag{A6}$$

This leads to

$$(\boldsymbol{R} - \lambda \boldsymbol{I})\boldsymbol{\beta} = \boldsymbol{0}. \tag{A7}$$

The eigenvalue $(\lambda_1, \ldots, \lambda_k, \ldots, \lambda_m, \ldots, \lambda_p)$ and respective eigenvalue to eigenvector can then be found

$$\begin{aligned} &\boldsymbol{B}_{p \times p} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k, \ldots, \boldsymbol{\beta}_m, \ldots, \boldsymbol{\beta}_p) = [\beta_{ik}]_{p \times p} \\ &\begin{cases} \beta_k' \beta_{k'} = 1 & k = k' \\ \beta_k' \beta_{k'} = 0 & k \neq k' \quad \text{independence.} \end{cases} \end{aligned} \tag{A8}$$

Therefore, $\boldsymbol{B}' \boldsymbol{B} = \boldsymbol{I}$.

According to (A6) $Var(y_k) = \boldsymbol{\beta}_k' \boldsymbol{R} \boldsymbol{\beta}_k = \lambda_k$ of component y_k; $k = 1, 2, \ldots, p$ can be obtained and $\boldsymbol{B}' \boldsymbol{R} \boldsymbol{B} = \boldsymbol{\Lambda}$. Alternately, based on (A6), $\boldsymbol{R}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{\Lambda} \rightarrow \boldsymbol{B}' \boldsymbol{R} \boldsymbol{B} = \boldsymbol{B}' \boldsymbol{B} \boldsymbol{\Lambda}$, $\boldsymbol{B}' \boldsymbol{R} \boldsymbol{B} = \boldsymbol{\Lambda}$ when $\boldsymbol{B}' \boldsymbol{B} = \boldsymbol{I}$, where the diagonalization matrix

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}. \tag{A9}$$

Finally, based on (A8) principle factor loading matrix $\boldsymbol{A}$, communality h_i, eigenvalue λ_k and contribution rate λ_k / p, accumulated contribution rate can be obtained (see Table A2)

$$\boldsymbol{A} = [a_{ik}] = [\sqrt{\lambda_k} \beta_{ik}]. \tag{A10}$$

In the real word, take $k = 1, 2, \ldots, m$ and $m < p$ to determine the accumulated contribution rate $\sum_{k=1}^{m} \lambda_k / p$.

APPENDIX-C: SELECTIVE RAW DATA ($C^k = 10$)

TABLE A.4

TABLE-A3 SELECTIVE RAW DATA ($C^k = 10$)

X	Y	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

APPENDIX-D: MORE INTERPRETATION OF TABLE 3

TABLE A.5

MORE INTERPRETATION OF TABLE-3

Selected Bitmap data	Pass/Fail	Event-1	Event-2	Event-3	Event-4	Event-5	Event-6	FM-1	FM-2	FM-3	FM-4	FM-5	FM-6	FM-7	FM-8	FM-9	FM-10	FM-11	FM-12	FM-13	FM-14	FM-15	FM-16	FM-17	FM-18	FM-19	FM-20
S-2	F	0	0	0	1	1	1	8	10	9	10	12	11	126	123	104	84	20	20	19	19	87	143	58	115	43	80

Die-2 (S-2); This die is a defective die (F).

This die is not influenced by event1~3, but by event4~6.

In bit map data of this die, there are 8 failures of failure mode –1 in this die (reminding, there could be several mega bits of SRAM in this die)

REFERENCES

[1] L. Yan, "A PCA-based PCM data analyzing method for diagnosing process failures," *IEEE Trans. Semiconduct. Manuf.*, vol. 19, no. 4, pp. 404–410, Nov. 2006.

[2] Z. Qian, F. Siegelin, B. Tippelt, and S. Muller, "Localization and physical analysis of a complex SRAM failure in 90 nm technology," *Microelectron. Reliab.*, vol. 46, nos. 9–11, pp. 1558–1562, 2006.

[3] C. Weber, "Yield learning and the sources of profitability in semiconductor manufacturing and process development," *IEEE Trans. Semiconduct. Manuf.*, vol. 17, no. 4, pp. 590–596, Nov. 2004.

[4] K. R. Skinner, D. C. Montgomery, G. C. Runger, J. W. Fowler, D. R. McCarville, T. Reed Rhoads, and J. D. Stanley, "Multivariate statistical methods for modeling and analysis of wafer probe test data," *IEEE Trans. Semiconduct. Manuf.*, vol. 15, no. 4, pp. 523–530, Nov. 2002.

[5] K. Imai and T. Kaga, "A novel filtering method to extract three critical yield loss components (gross, repeated, and random) FIMER," *IEEE Trans. Semiconduct. Manuf.*, vol. 13, no. 4, pp. 408–415, Nov. 2000.

[6] J. Horan and C. Lyden, "Rapid IC performance yield and distribution prediction using a rotation of the circuit parameter principals components," *Microelectron. Reliab.*, vol. 38, no. 12, pp. 1913–1918, Dec. 1998.

[7] D. A. White and D. Boning, "Spatial characterization of wafer state using principal component analysis of optical emission spectra in plasma etch," *IEEE Trans. Semiconduct. Manuf.*, vol. 10, no. 1, pp. 52–60, Feb. 1997.

[8] T. Hamada and M. Sugimoto, "Application of a bitmap analysis system to the forefront of DRAM devices development," in *Proc. IEEE/SEMI Adv. Semiconductor Manuf. Conf.*, 1997, pp. 222–227.

[9] R. S. Collica, J. P. Card, and W. Martin, "SRAM bitmap shape recognition and sorting using neural networks," *IEEE Trans. Semiconduct. Manuf.*, vol. 8, no. 3, pp. 326–332, Aug. 1995.

[10] I. T. Jollife, *Principal Component Analysis*. New York: Springer-Verlag, 2002, ch. 1.

[11] D. Y. Shee and G.-H. Tzeng, "The key dimensions of criteria for the evaluation of ISPs: An exploratory study," *J. Comput. Inform. Syst.*, vol. 42, no. 4, pp. 112–121, 2002.

Yeou-Lang Hsieh was born in Taiwan in 1973. He received the Bachelors degree in engineering science from National Cheng Kung University, Tainan City, Taiwan, in 1995, the Masters degree in electronic engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1997, and the second Masters degree in management of technology from NCTU in 2005. He is currently pursuing the Ph.D. degree from the Institute of Management of Technology, NCTU.

He joined Taiwan Semiconductor Manufacturing Company (TSMC) as a Process Integration Engineer in 2000, and since 2004 has been a Principal Product Engineer. In this role, he took part in the Advanced Technology Yield Improvement Team. He is currently a Project Manager with the Customer Technical Supporting Division. During his position with TSMC, he has registered six U.S. patents for TSMC.

Gwo-Hshiung Tzeng (F'02) was born in Taiwan in 1943. He received the Bachelors degree in business management from the Tatung Institute of Technology, Taipei, Taiwan, in 1967, the Masters degree in urban planning from Chung Hsing University, Taichung, China, in 1971, and the Ph.D. degree in management science from Osaka University, Osaka, Japan, in 1977.

He was an Associate Professor with National Chiao Tung University, Hsinchu, Taiwan, from 1977 to 1981, a Research Associate with the Argonne National Laboratory, Argonne, IL, from July 1981 to January 1982, a Visiting Professor with the Department of Civil Engineering, University of Maryland, College Park, from August 1989 to August 1990, a Visiting Professor with the Department of Engineering and Economic System, Energy Modeling Forum, Stanford University, Stanford, CA, from August 1997 to August 1998, a Professor with Chaio Tung University from 1981 to 2003, and is currently a Chair Professor with the Institute of Management of Technology, National Chiao Tung University. His current research interests include statistics, multivariate analysis, network routing and scheduling, multiple criteria decision making, fuzzy theory, hierarchical structure analysis for applying to technology management, energy, environment, transportation systems, transportation investment, logistics, location, urban planning, tourism, technology management, electronic commerce, global supply chain, and so on.

Dr. Tzeng received the National Distinguished Chair Professor (highest honor offered by the Ministry of Education Affairs, Taiwan) and Distinguished Research Fellow (highest honor offered by NSC, Taiwan) in 2000. He got the Highly Cited Paper (March 13, 2009) ESI for "Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS" as published in the *European Journal of Operational Research* in 2004, which has been recently identified by Thomson Reuters' Essential Science IndicatorsSM to be one of the most cited papers in the field of economics. He received the Pinnacle of Achievement Award of the World in 2005, the National Distinguished Chair Professor and Award (highest honor offered) from the Ministry of Education Affairs of Taiwan, the Distinguished Research Award thrice, and the Distinguished Research Fellow (highest honor offered) from the National Science Council of Taiwan twice. He organized a Taiwan Affiliate Chapter of the International Association of Energy Economics in 1984 and was the Chairman of the Tenth International Conference on Multiple Criteria Decision Making, Taipei, from July 19 to 24, 1992. He was the Co-Chairman of the 36th International Conference on Computers and Industrial Engineering, Taipei, from June 20 to 23, 2006. He was the Chairman of the International Summer School on Multiple Criteria Decision Making, Kainan University, Taiwan, from July 2 to 14, 2006. He is a member of IAEE, ISMCDM, World Transport, the Operations Research Society of Japan, the Society of Instrument and Control Engineers of Japan, the City Planning Institute of Japan, the Behaviormetric Society of Japan, the Japan Society for Fuzzy Theory and Systems, and is participating in many societies of Taiwan. He is the Editor-in-Chief of the *International Journal of Operations Research*, the *International Journal of Information Systems for Logistics and Management*, and others.

Grace Tyng-Ruu Lin received the Ph.D. degree from the Judge Business School of Cambridge University, Cambridge, U.K., in 2003.

She is currently an Associate Professor with the Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan. She has been published in the *Journal of Business Ethics*, *Technology in Society*, *Developing Economies*, *Entrepreneurship and Regional Development*, the *Journal of Scientific and Industrial Research*, the *Journal of Global Business and Technology*, *Technovation*, and others. Her current research interests include innovation and technology management, strategic management, and marketing studies.

Hsiao-Cheng Yu received the Ph.D. degree from the School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, in 1981.

He is currently a Professor with the Institute of Management of Technology, National Chiao Tung University (NCTU), Hsinchu, Taiwan. Before joining NCTU, he was a Member of Technical Staff, AT&T Bell Laboratories, Murray Hill, NJ, from 1985 to 1992, and was a Consultant with Contel Information Systems, Great Neck, NY, from 1981 to 1985. His current research interests include telecommunications and broadcasting policies, entrepreneurship and venture capital, telecommunications technologies and service management, and trade skills and business paradigms in high-tech industries.