Project Categorization and Assessment Utilizing Multivariate Statistical Techniques to Facilitate Project Pattern Recognition, Categorization, Assessment and Pattern Migration Over Time

By Bob Prieto
Fluor

Disclosure: One or more of the concepts, techniques or methodologies described in this paper are protected by various patents and/or service marks pending. Owners and clients interested in discussing any of these concepts further should contact the author.

Large scale projects, especially capital construction projects, are notoriously difficult to manage. Project managers, or other stakeholders, require techniques to quickly assess a current state of their project and to better anticipate likely performance trajectories. Ideally, project managers would be able to quickly compare their project against historical projects or known best practices in order to benefit from the wealth of prior experience that exists. Unfortunately, there are very few techniques currently available to project managers that allow them to recognize the project's state as being similar to circumstances related to other projects.

The Opportunity

This paper focuses on understanding, categorizing, assessing and monitoring a project based on consideration of a large number of project values in order to create a pattern definition analogous to a "picture" of the project. The concept described herein is the subject of a patent filing and as such enjoys all the protections of that process.

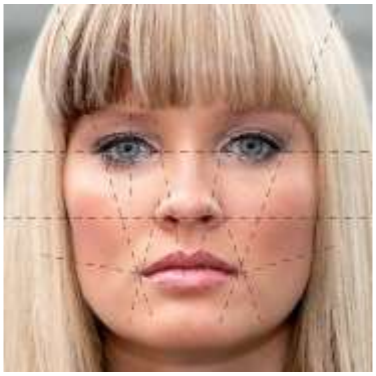
The picture created through consideration of a large number of project values is then compared with other similar pictures and like pictures are similarly grouped and categorized where each category has certain common descriptive features and by extension some reasonably anticipated project attributes. This project picture may be retaken over time (the project lifetime, potentially including its operating phase) and its strength of correlation with its initially assigned group measured (is it more or less similar to its original group).

Changed Characterization

Over time it is possible for a project "picture" to suggest that the project should be otherwise categorized. This may result from one of two circumstances.

The first would be a significant enough change to the project "picture" over time such that it no longer ideally fits in the originally assigned group. Such reclassification would suggest that the project has different common descriptive features and attributes and suggests changed areas of management focus and attention and new project areas of interest. This migration of project type allows the project manager to understand the nature of change that the project is experiencing and to seek out insights from better fitting project archetypes.

The second instance which might drive project reclassification would be the result of changes in the composite library of all project pictures such that groupings or the definitions of their characteristics changed as sample size grew. Such changes in the composite library, ultimately, will allow for more precision in project characterization at early project stages including at the conception and initial project planning stages when various execution options may be under consideration.



An analogy would be in the area of facial recognition where multivariate statistical techniques are used. In this analogy we may take 10 pictures of an individual over time and we would expect the 10 images to statistically group. If we took similar sets of photos over time of other individuals we would expect similar images to group. Examples might be Caucasian men, Southeast Asian women, Eskimo women and the such. If one individual developed a facial scar or has cosmetic surgery it is possible that an individual might be subsequently assigned to a different facial recognition group.

This analogy is an apt description of the application of multivariate statistical techniques to identifying and characterizing project characteristics in order to better enable the project manager to manage these endeavors. Previously there has been a more limited recognition of using multivariate analysis to identify interrelationships among project factors and financial performance but as envisioned here these insights would allow us to anticipate challenges and influence outcomes at an earlier stage.

Helping the PM Focus

Projects are complex affairs with a wide range of project values. These may include values with respect to the client, his project manager, contract form; project values such as project type, size, location, project challenges; and execution contractor values such as risk assessment, project manager, execution approach, work load; stakeholder values such as number, degree of support, prior history; and a range of other project values.

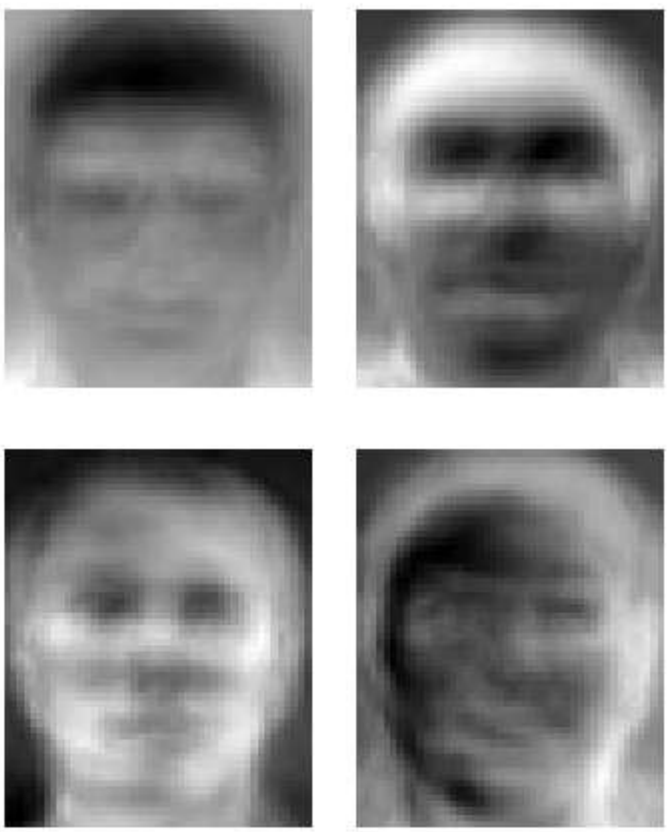
Assessing which areas to focus on; areas of likely challenge; opportunity areas to explore and so on are not always readily evident. Utilizing a pattern recognition type approach based on multivariate statistical techniques provides the project manager with an additional tool to manage the project.

This pattern recognition type approach also provides a basis for identifying subtle but pervasive changes to the project picture from potentially correlated common drivers. Think of the project picture as either darkening or lightening. Early recognition of potential common drivers acting on the project provides an ability to seek out, understand and manage these drivers to the advantage of the project. Examples could include constraint coupled factors or risks not readily apparent or systemic factors and risks with complex inter-relationships.

This pattern recognition approach utilizing multivariate analysis would also support a more granular categorization of discrete aspects of the project. As an example a "picture" of stakeholder relationships taken over time would not only allow early categorization facilitating strategy identification but also identifying more subtle shifts and trends. In this respect, statistically meaningful portions of the "picture" might be compared for more granular characterization or analysis. This would be similar to looking at the eyes of all Caucasian men to further categorize by shape or color.

The application of the proposed approach would similarly be applicable on a program or project portfolio basis.

Creating the Fluor Eigenproject[SM]



A data set initially composed of project "pictures" from different points in time from completed projects provides a basis for initial "group" definition. These definitions will be strengthened as additional pictures are subsequently added. In effect the initial data set acts as a training set for the final developed tool.

"Pictures" from the same project over time are stacked to create a multivariate image where time in effect is the third dimension and each individual image is comprised of rows and columns of "pixels" where columns correspond to similar types of project value (client; complexity; environmental factors; stakeholder).

These multivariate images are used to confirm group classification while individual pictures provide some initial sense of directionality (classification is getting stronger or weaker). Within a given group, the anticipated changes in the multivariate picture over time can be modeled and provides another tool to assess project evolution. Select defined portions of the picture (analogous to eyes, mouth, ears, nose etc) may be separately characterized to support more granular analysis but trading off some of the insights into non obvious correlations.

"Pixels" are considered to be the variables in the analysis and for a given project these variables are expected to be highly correlated. Initially these variables are assumed to be those variables routinely tracked as part of a project management system and reported on project status reports.

For example, pictures comprised of all values (pixels) from a complete Project Status Report at each period through the project lifetime are used to create the initial project image database and categorization.

The developed tools "unpack" each image column wise such that sequence of parameters would be the same for each project picture so unpacked. This would allow for subsequent analysis of select portions of the data set such as stakeholder data; productivity related data; and external factors assessments.

Each picture's data can be thought of as another row of unpacked data. The database of unpacked data is cumulative such that as a new picture is added to the database, effectively one additional row is added. The data matrix would be equal to the number of "pixels" x the number of "pictures".

The first step in utilization of the techniques described is classification of each given project "picture" into the defined groups. This classification will be independently undertaken utilizing two different but related multivariate statistical techniques.

The first technique utilized in classification is linear discriminant analysis or LDA and it is a statistical technique employed in pattern recognition such as facial or voice recognition. LDA classifies patterns based on the so-called Mahalanobis distance. In statistics, Mahalanobis distance is a distance measure introduced by P. C. Mahalanobis in 1936. It is based on correlations between variables by which different patterns can be identified and analyzed. It gauges similarity of an unknown sample set to a known one. It differs from Euclidean distance in that it takes into account the correlations of the data set and is scale-invariant. In other words, it is a multivariate effect size.

An effect size calculated from data is a descriptive statistic that conveys the estimated magnitude of a relationship without making any statement about whether the apparent relationship in the data reflects a true relationship in the population.

Each project "picture" is classified into the group whose mean is closest to it in the Mahalanobis sense.

LDA relies on key assumptions with respect to normal distribution of multivariate conditional probabilities and equivalence of so called group covariance matrices. These assumptions allow simplification of the analysis and are useful in all but truly first of a kind projects where correlation with any defined group is weak at best. In conducting LDA a common group covariance matrix would be required but to strengthen overall analysis in this type of pattern recognition a pooled covariance matrix of all the groups in the project database would be used instead.

The covariance matrix will be exceptionally large and not feasible to estimate. Techniques such as using PCA (principal component analysis) to reduce dimensionality by extracting so called principal components is equally infeasible.

The data matrix however is equal to the number of "pixels" x the number of "pictures" and we may think of the associated covariance matrix consisting of a number of non zero eigenvectors equal to the number of project pictures in the data set.

From this set of calculated eigenvectors we may utilize PCA to extract an analogue to eigenfaces (facial recognition) or eigenvoices (speech recognition) which we have called eigenprojects. Any given project "picture" can be reconstructed by projecting it onto the eigenprojects with reconstruction complete when it has been projected using all the eigenprojects.

For practical purposes, one can project images only onto approximately the first 20 eigenprojects and these new variables (principal component scores) used in LDA such that the data matrix is equal to number of project pictures x 20 eigenprojects. Sensitivity test of the training data sets will confirm the appropriateness of limiting projection to 20 eigenprojects by calculating the "Apparent Error Rate (APER)" which is an optimistic assessment of the actual error rate.
The second technique utilized in classification is the Fisher Discriminant method which is similar in objective to PCA in the sense that it seeks to reduce dimensionality. It does not make some of the assumptions of LDA such as normally distributed classes or equal class covariances. Utilization of two dimensional Fisher discriminant space plots is a useful tool to visualize the proximity of various groups in the classification system.

The second step in utilization of this technique is assessing which areas that the project manager should focus on given the similarity of his project to some respective group of projects for which we have developed some insight. Insights for the initial eigenprojects database come from a review of contemporaneously prepared project reports, lessons learned reports prepared and selective interviews with project managers and executives to capture deeper insights.

Realizing the Opportunities

Assessment of initial database projects and other subsequent projects captured in the ever growing project "picture" database allows the PM to identify areas of likely challenge; opportunity areas to explore and suggests other factors based on pattern relevant experience that are not otherwise readily evident. Utilizing a pattern recognition type approach based on multivariate statistical techniques provides the project manager with an additional tool to manage the project.

A third step, an opportunity, is in utilization of this technique as a basis for identifying subtle but pervasive changes to the project picture from potentially correlated common drivers. Think of the project picture as either darkening or lightening.


Early recognition of potential common drivers acting on the project provides an ability to seek out, understand and manage these drivers to the advantage of the project. In this case the comparative analysis would be between project "pictures" taken at different times. Both absolute comparison and comparison of respective eigenproject values from the database would be accomplished. Additionally, it is possible to compare eigenproject movies of a current project with a similar eigenproject movie for the correlated group allowing a deeper understanding of how group values change with "aging".

This later capability to understand project aging patterns would have particular relevance in the operating and maintenance phase of the project.

A fourth step, also an opportunity area, is in utilization of this technique as a basis for identifying subtle but pervasive changes in sub-elements of a project. Caution is required to assure meaningful conclusions can be drawn. In particular, there may be relevance in evaluating complex stakeholder situations or assumption migration in complex, long duration projects.

Finally, this technique can be applied across a program but in this instance the training database would be required to be programmatic in nature.

The application of multivariate analysis techniques, such as those used in facial or voice recognition, to project “pictures” allows us to see patterns that might not be otherwise easily recognizable and to categorize projects in ways that allow project managers to draw upon a broad array of project experience and lessons learned, focusing on those most similar to the project “picture” at hand.

As various project execution approaches are considered, alternative “pictures” may be assessed and compared to similar composite pictures in an ever expanding database of such insights.

Value Creation

Specifically, the opportunity exists to:

- Create a reference class of project characterizations or eigenprojects.
- Utilize a statistical methodology to categorize projects into a multivariate correlated group.
- Link reference class projects to a database of project management characteristics, including, but not limited to risk classes and types, recurring project issues; potential opportunities; project parameters to be monitored and other such factors to help support management of a project with many attributes.
- Provide a methodology for detecting potential systemically linked changes broadly affecting a project.
- Create a time phased reference class project profile to correlate and predict potential project correlations over time. A so called eigenproject movie
- Apply this methodology to significant sub-elements of a project such as stakeholder management, assumption tracking and correlation and productivity impacting factors.
- Alternatively, separate eigenproject pictures may be created along each of the three bottom lines comprising the so called "triple bottom line"
- Manage portfolios of projects or other such portfolios characterized by large numbers of non obviously correlated variables where a portfolio "picture" database replaces the contemplated project database.
- Apply these concepts during the operating phase of a project against an operating phase eigenproject database to detect subtle changes in plant operating condition and operational readiness and provide predictive insight into future maintenance areas of focus.

Currently, project characterizations do not capture subtle or perhaps indirect relationships between project attributes which are masked by project complexity. The approach outlined in this paper will see relationships without necessarily needing to understand what the common links or drivers are.

Other Value Creating Applications

Bonding and reinsurance companies serving the construction industry would benefit from being able to categorize projects with a high degree of complexity based on multivariate pattern recognition.

Project managers and PMO organizations in other industries would similarly benefit from such an ability to group complex projects and monitor and anticipate their change through the project execution period.

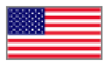
References

1. Generalized Analysis of Value Behavior over Time as a Project Performance Predictor; PM World Journal; Volume I Issue III; October 2012; http://pmworldjournal.net/?article=2442
2. “Generalized Analysis of Value Behavior Over Time as a Project Performance Predictor”; College of Performance Management (pending)
3. Project Categorization And Assessment Through Multivariate Analysis; Attorney Docket 100325 0516PRO
4. Time Derivative-Based Program Management Systems And Methods; Attorney Docket 100325-0501US
5. Multi-Dimensional Life Cycle Project Execution System; Attorney Docket No. 100325.0464US1
6. Life Cycle Analysis; Attorney Docket 100325-0527PRO
7. The GIGA Factor; Program Management in the Engineering & Construction Industry; CMAA; ISBN 978-1-938014-99-4; 2011

About the Author

Bob Prieto

Senior Vice President
Fluor



Bob Prieto is a senior vice president of Fluor, one of the largest, publicly traded engineering and construction companies in the world. He is responsible for strategy for the firm's Industrial & Infrastructure group which focuses on the development and delivery of large, complex projects worldwide. The group encompasses three major business lines including Infrastructure, with an emphasis on Public Private Partnerships; Mining; and Industrial Services. Bob consults with owner's of large engineering & construction capital construction programs across all market sectors in the development of programmatic delivery strategies encompassing planning, engineering, procurement, construction and financing. He is author of "Strategic Program Management" and "The Giga Factor: Program Management in the Engineering and Construction Industry" published by the Construction Management Association of America (CMAA) and "Topics in Strategic Program Management" as well as over 450 other papers and presentations.

Bob is a member of the ASCE Industry Leaders Council, National Academy of Construction and a Fellow of the Construction Management Association of America. Bob served until 2006 as one of three U.S. presidential appointees to the Asia Pacific Economic Cooperation (APEC) Business Advisory Council (ABAC), working with U.S. and Asia-Pacific business leaders to shape the framework for trade and economic growth and had previously served as both as Chairman of the Engineering and Construction Governors of the World Economic Forum and co-chair of the infrastructure task force formed after September 11th by the New York City Chamber of Commerce.

Previously, he established a 20-year record of building and sustaining global revenue and earnings growth as Chairman at Parsons Brinckerhoff (PB), one of the world's leading engineering companies. Bob Prieto can be contacted at Bob.Prieto@fluor.com.