Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

CLIMA 2016 - proceedings of the 12th REHVA World Congress

volume 6

Heiselberg, Per Kvols

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Heiselberg, P. K. (Ed.) (2016). *CLIMA 2016 - proceedings of the 12th REHVA World Congress: volume 6*. Department of Civil Engineering, Aalborg University.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: May 15, 2024

Applicability of statistical shape analysis for estimating mean shape of building groups

DeukWoo Kim[#1], Jiyoun Lim[*2], Seungeon Lee[*3]

[#] *Convergence Research Planning Division, Korea Institute of Civil Engineering and Building Technology, Goyang, 411-712, South Korea*

[1] deukwookim@kict.re.kr

[*] *Building and Urban Research Institute, Korea Institute of Civil Engineering and Building Technology, Goyang, 411-712, South Korea*

[2] jiyoun@kict.re.kr

[3] selee2@kict.re.kr

Abstract

Model-based scenario analysis is an effective way to underpin revision of energy codes/standards and assess new technologies rationally. Many studies have been conducted to develop standard building energy models, e.g. reference buildings, and secure the model validity. The reference building is created based on national statistics on building stocks in general. However, the shape of buildings is typically overlooked and thus the shape is decided on the basis of an expert's opinion. One of the main reasons for this is the lack of a well-defined framework for handling the building shape in a quantifiable way. The building shape is closely related to cooling and heating loads due to the impact of solar radiation and airflows. Hence in order to strengthen the foundation of the reference model framework, it is necessary to address how to create a representative building shape from a building stock.

In this paper, the applicability of statistical shape analysis, which is an analysis of the geometrical properties of some given set of shapes by statistical methods, is investigated to estimate representative building shapes and their variations. The coordinate data of building footprints were obtained by Geographic Information Service of the Statistics Korea, and analyzed by statistical shape analysis. Finally, a mean shape of a building stock was derived as proof of concept. The established framework and the usefulness of the results are described in detail.

Keywords – reference building; building shape; building footprints; statistical shape analysis; GIS

1. Introduction

Model-based performance evaluation can be effectively used to predict energy savings according to revision of energy codes/standards or technical improvements (insulation, high performance window, HVAC etc.) (Deru et al. 2011, IEE 2012, European Commission 2012). When using the model, it is important to prepare statistically meaningful input values (i.e. reference values) for reliable simulation results

(Waltz, 2000). Hence the framework for developing reference buildings has been vigorously discussed (Corgnati et al. 2013).

In the case of U.S.A., the reference model design guidelines were proposed by aggregating existing case studies (Hendron and Engebrecht 2009, Deru et al. 2011). The developed reference models were widely used to underpin the revision of energy-related codes/standards or industrial applications. Likewise, in Europe (13 countries), the typology of residential buildings was discussed and classified based on national statistics (IEE 2010, IEE 2012). Each country involved in this undertaking has expended efforts to establish representative building models (including input values) to obtain reliable results. Despite these efforts, there are still loose ends regarding how the decision on *representative building shapes* were derived.

According to previous studies (Deru et al. 2011, IEE 2009, IEE 2012, DOE 2015), representative building shapes are determined on the basis of expert experience (Ballarini et al. 2011). However, the grounds for such determination have been insufficiently explained, and the decision process is not clear. Anyone can raise a question with regard to the representativeness of shapes. Such a lack of discussion is presumably due to the fact that the shape itself is deemed to be qualitative information. For this reason, the reference shape is typically determined in consideration of height, gross floor area, aspect ratio, etc. (building metadata).

Recently, with the latest expansion of geographic information services (GIS) and LiDAR scanning technology, advanced map services have become available. In South Korea, the database on three-dimensional building shapes has been steadily built up (SPACEN, 2015). Thanks to this, it has become easy to obtain architectural shape information (e.g. coordinate data) on a nationwide scale. As a pilot study, the coordinate data of building footprints of a target area obtained from Statistics Korea were treated firstly. Finally, the mean building shape and its variation were derived by means of statistical shape analysis. The research process is explained in a later section in detail.

2. Literature review

2.1 Determination of the reference building model

A reference building is a theoretical model that is created based on national statistics and previous research findings. It is widely used to assess performance improvements, or to predict energy savings of advanced technologies quantitatively (Tommerup and Svendsen 2006, Gaglia et al. 2007, Kurnitskia et al. 2007, Desideri et al. 2012, Olofsson and Mahlia 2012). Moreover it can be used to evaluate achievable reductions in energy consumption under a certain energy policy applied to some type of buildings on a nationwide scale (Deru et al. 2011, European Commission 2012).

According to the European Commission (2012), the reference building is a building characterized by the representativeness of its functionality and geographic location, including indoor and outdoor climate conditions. It aims to represent the typical and average building stocks in terms of climatic conditions and functionality. The goal of the application of the reference building is to characterize the energy performance of typical building use under typical operations. Hence, it is necessary to reflect as accurately as possible the actual behaviour of the national building stock and to ensure the results are as representative as possible.

According to Corgnati et al. (2013), four sub-sets of features (operation, form, system, and envelope) should be gathered to create a reference building. Based on these four features, three types of reference buildings can be made; *Example Building, Real Building, Theoretical Building*. An *Example Building* is used when no statistical data are available, and it thus relies on the assumptions of experts and previous studies. A *Real Building* is the most typical building in a certain category with average characteristics based on statistical data; for example, mean conditioned area, mean U-value of opaque components, etc. It is therefore necessary to have a large amount of information on the building stock to create a *Real Building*. A *Theoretical Building* is a fictional building, which is mainly based on synthetic data in order to define a reference building as a statistical composite of the features found within a category of buildings in the stock. The *Theoretical Building* is therefore made of the most commonly used materials and systems. In a similar way, the U.S. Department of Energy developed commercial reference buildings (DOE 2015). These reference buildings consist of 16 building types that represent approximately 70% of the commercial buildings in the U.S. according to the report published by the National Renewable Energy Laboratory titled U.S. Department of Energy Commercial Reference Building Models of the National Building Stock (Deru et al. 2011).

These well-established outcomes of previous works (Deru et al. 2011, European Commission 2012, Corgnati et al. 2013) have been widely applied; however, these overlook how to derive representative building shapes from a building stock considering actual shapes. In most cases, the shapes of the reference building were determined on the basis of expert experience or estimated on the basis of indirect information e.g. height, aspect ratio, gross floor area, building area, etc., which are irrelevant to the actual shape.

2.2 Application of statistical shape analysis

Statistical shape analysis is an analysis on the geometrical properties of some given set of shapes by statistical methods. It has various applications, including medical imaging, biology, genetics, archeology, computer vision, machine learning, etc. (Dryden and Mardia 1998). One way to describe a shape is using a finite number of points on the outline (referred to as a *landmark*). A mathematical representation of a rectangular shape in 2-dimensional space can be expressed as $\mathrm{X} = [x_1, x_2, x_3, x_4; y_1, y_2, y_3, y_4]$. Important aspects of the shape analysis are to obtain a measure of distance between shapes, to estimate mean shapes from samples, to estimate shape variability, to perform clustering, and to test for differences between shapes (Dryden and Mardia 1998).

All shapes need to be aligned properly to measure their similarity. An often used method is the Procrustes analysis, which filters out the translation, scale, and rotation effect of an object (Dryden and Mardia, 1998). A classic example is to quantify differences between male and female Gorilla skull shapes, normal and pathological bone shapes, etc. Fig. 1 shows hierarchies of the shape space depending on the removal of translation, scale, rotation. The Procrustes analysis brought the set of shapes into a frame of reference, in which we can analyse and compare various building shapes in a quantitative way.

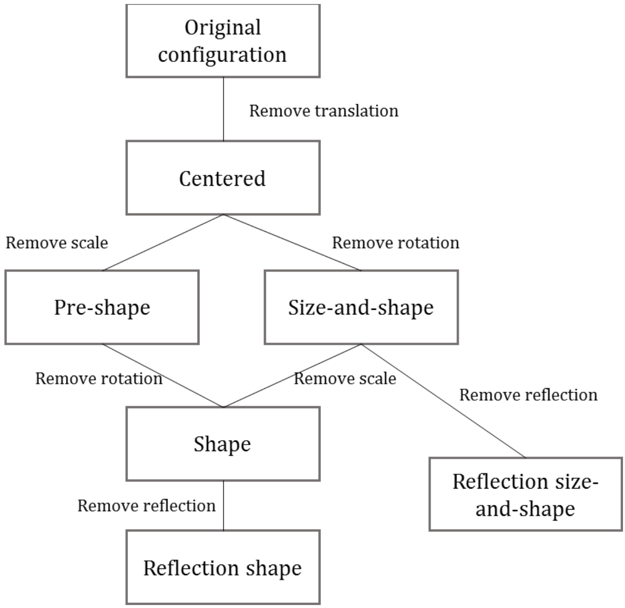


Fig. 1. Hierarchies of various shape spaces (Dryden and Mardia 1998). The reflection shape is all the geometrical information that is invariant under similarity transformations and reflection.

3. Methodology

3.1 Statistical shape analysis

The Procrustes analysis uses a shape metric, the so-called Procrustes distance, to measure similarity between objects. The Procrustes distance is a least-squares type shape metric that requires two aligned shapes with one-to-one point correspondence. The alignment involves four steps; (i) computation of the centroid of each shape, (ii) re-scale of each shape to have equal size, (iii) alignment with respect to position two shapes at their centroids, (iv) alignment with respect to orientation by rotation (Stegmann and Gomez 2002). The result of the steps is a set of shapes which is superimposed with normalization. Based on this process, a mean shape, whose coordinates are the average of that of the aligned shapes, can be easily calculated. If applied to a group of buildings, a *mean* building shape can be obtained.

Meanwhile, assuming that a number of building footprints overlap, the image may be seen as a set of random shapes. Aside from calculating the mean building shape, it would be interesting to estimate a *variation* of the overlapped shapes. Thanks to principal component analysis (PCA), it is possible to extract the principal components in shape variations. PCA is a statistical procedure that uses an orthogonal transformation to convert a set of data of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. In other words, it estimates the orthogonal vectors (namely, principal components) that best describe the variance of a given data set. In summary, a mean building shape and its variance can be estimated applying the Procrustes analysis and PCA techniques.

3.2 Data collection

The coordinate data (*.shp file extension) of building footprints are obtained from Geographic Information Service of the Statistics Korea for statistical shape analysis. The *.shp is a format that contains a geospatial vector data. It is developed and regulated by Environmental Systems Research Institute (ESRI) as an open specification for data interoperability among GIS software products. The shapefile stores non-topological geometry information and information about attributes for the spatial features in a data set. Hence it can spatially describe vector features: points, lines, and polygons, representing, for example, building footprints, rivers, and roads (ESRI 1998).

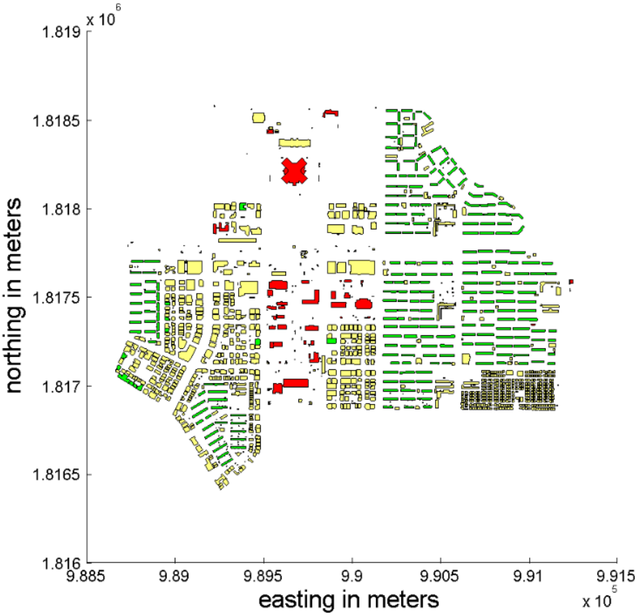
4. Results

Dunsan-dong, in Daejeon, South Korea was selected as the target area (Fig. 2). By reference to principal building use codes (CLASS_CD, Table 1) in the shapefile, the coordinate data of the building footprints were selected. In this study, apartment buildings (CLASS_CD: 012001) and public/administrative buildings (CLASS_CD: 10207) were selected as shown in Fig 2(b) to give a demonstration. The number of the apartment buildings and the public/administrative buildings were 254, and 41 respectively (total number of the buildings is 1476). In South Korea, the principal use code (CLASS_CD, Table 1) is classified into 49 items such as residential, public/administrative, educational, medical, religious, accommodation etc. The coordinate data of the building footprints were obtained by reference to the GEOM attribute (Table 1), which contains the footprints as multi-polygon style. In this paper, only quadrilateral-shape buildings were analyzed by applying the Procrustes analysis for ease of use.

Table 1. Database schema of building attributes

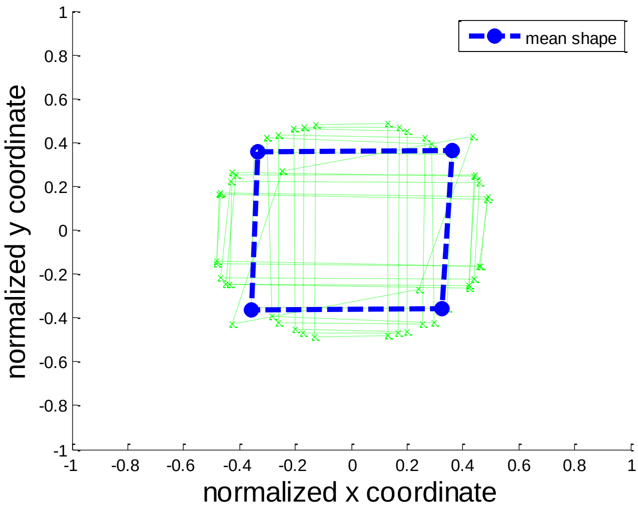
Attribute	Description	Data type	PK/FK
OBJECTID	unique id	number	pk
ADM_DR_CD	province code (type 1)	varchar2	fk
EMD_CD	province code (type 2)	varchar2	fk
SIG_CD	city code	varchar2	
RD_NM	street code	varchar2	
BULD_MNNM	main building code	varchar2	
BULD_SLNO	second building code	varchar2	
BULD_NM	registered building name	varchar2	
CLASS_CD	principal use code	varchar2	fk
GEOM	building shape coordination	multi-polygon	

(a) A screenshot of target area: Dunsan-dong in Daejeon City, South Korea (red line)

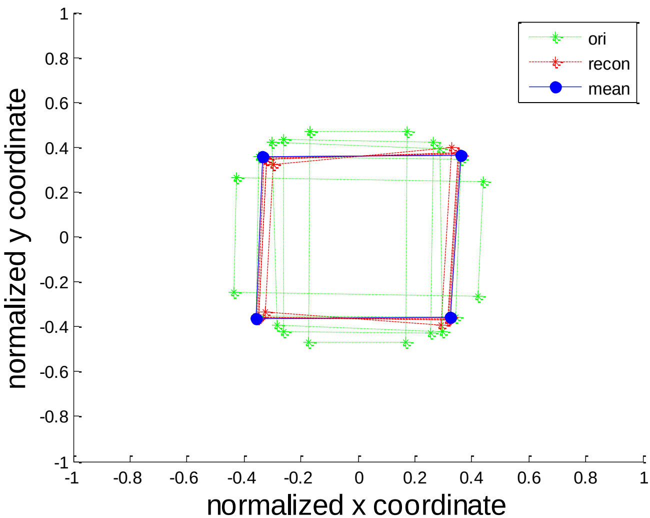


(b) Building footprints in shapefiles: the red and green colors denote public/administrative agency and apartment buildings respectively

Fig. 2. Target area and building shape information

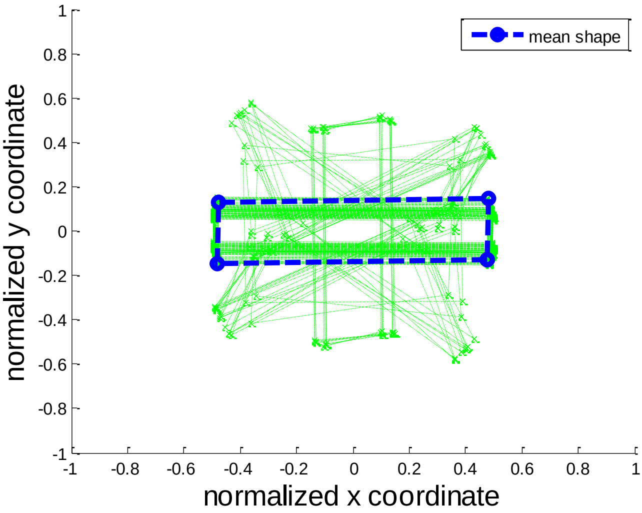


(a) The estimated mean shape (blue-line) from the public/administrative agency buildings (green line)

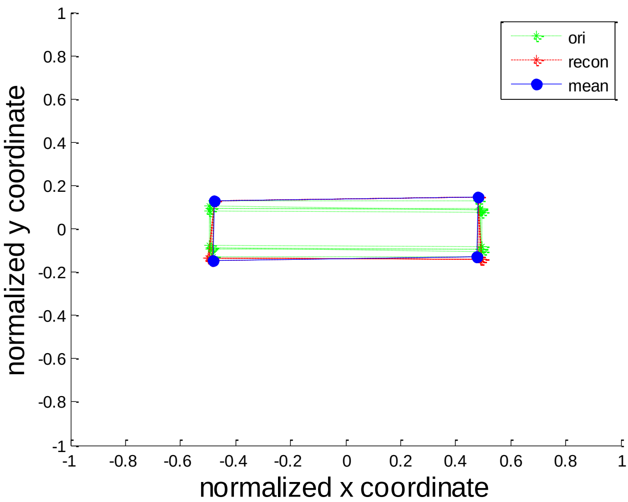


(b) The approximated five shapes (red asterisk) from the original shpaes using PCA

Fig. 3. Results of the Procrustes analysis on the public/administrative agency buildings



(a) The estimated mean shape (blue-line) from the apartment buildings (green line)



(b) The approximated five shapes (red asterisk) from the original shpaes using PCA

Fig. 4. Results of the Procrustes analysis on the apartment buildings

Figs. 3 and 4 show the results of the Procrustes analysis regarding public/administrative agency buildings and the apartment buildings respectively. All shapes are aligned near (0, 0) coordinates and normalized from -1 to 1 level (Figs. 3(a) and 4(a)). The blue dotted line with a circle (Quadrilateral) is the *mean* of the sample footprints (dotted green line with a cross). It is shown that the clustered vertices of the sample footprints located near the mean shape are more clearly observed due to the alignment and normalization.

It is possible to estimate *principal directions* of the sample vertices (covariance of the vertices between the sample and the mean) using the principal component analysis. A certain shape instance can be effectively approximated by multiplying outputs of the PCA (e.g. eigenvector and eigenvalue) based on the mean shape. Figs. 3(b) and 4(b) show the reconstructed five shapes (red-dashed line) from the original data (green-dotted line). The result of Fig. 4(b) is better than that of Fig. 3(b); however, the accuracy of the reconstruction is beyond the scope of this study and will be examined in subsequent studies.

5. Conclusion

Building shapes are typically overlooked and thus are usually determined on the basis of expert opinion. One of the main reasons for this is the lack of a well-defined framework for the handling of building shapes (images) in a quantifiable way. This paper has demonstrated that statistical shape analysis technique can be effectively applied in the creation of a reference building. Firstly, the shapefile was obtained from Geographic Information Service of the Statistics Korea, and the geometrical information (building footprints) was extracted for statistical shape analysis. Finally, the Procrustes analysis was applied, and the mean shape and variations of the building groups (apartment buildings and public/administrative buildings) were derived as proof of the concept. In order to strengthen the foundation of the reference model framework, our suggested approach can fill the gap to create a representative building shape from a building stock.

Acknowledgment

This research was supported by a grant (15AUDP-B079104-02) from Architecture & Urban Development Research Program funded by the Ministry of Land, Infrastructure and Transport of Korean government

References

[1] SPACEN (2015), V-World, Spatial Information Industry Promotion Institute, http://data.vworld.kr/data/dc_datinfo_s001.do (accessed 2015.12.11.)

[2] Corgnati, S.P., Fabrizio, E., Filippi, M., Monetti, V. (2013), Reference buildings for cost optimal anlaysis : Method of definition and application, Applied Energy, 102, 983-993.

[3] Deru, M., Field, K., Studer, D., Liu, B., Halverson, M., Winiarski, D., Rosenberg, M., Yazdanian, M., Huang, J., Crawley, D., (2011), U.S. department of energy commercial reference building models of the national building stock, Technical report, National Renewable Energy Laboratory, 2.

[4] DOE (2015), Commercial reference buildings, Available at http://energy.gov/eere/buildings/commercial-reference-buildings (Last accessed 2015.10.26.)

[5] Dryden, I.L., Mardia, K.V. (1998), Statistical Shape Analysis, John Wiley & Sons, Inc.Waltz, J.P., (2000), Computerized Building Energy Simulation Handbook, Fairmont Press
[6] EIA. (2005). 2003 Commercial Buildings Energy Consumption Survey. Washington, DC: Energy Information Administration. Available at http://www.eia.gov/consumption/commercial/data/2003/ (Last accessed 2015.10.26.).
[7] ESRI (1998), ESRI Shapefile Technical Description: An ESRI White Paper, Environmental Systems Research Institute, Inc.
[8] European Commission (2012), Guidelines accompanying Commission Delegated Regulation (EU) No 244/2012 of 16 January 2012 supplementing Directive 2010/31/EU of the European Parliament and of the Council on the energy performance of buildings by establishing a comparative methodology framework for calculating cost-optimal levels of minimum energy performance requirements for buildings and building elements, Official Journal of the European Union
[9] Gaglia, A.G., Balaras, C.A., Mirasgedis, S., Georgopoulou, E., Sarafidis, Y., Lalas, D.P. (2007) Empirical assessment of the hellenic non-residential building stock. Energy consumption, emissions and potential energy savings, Energy Conversion and Management, 48(4), 1160-1175
[10] Hendron, R. & Engebrecht, C. (2009), Building america research benchmark definition, Technical report, National Renewable Energy Laboratory
[11] IEE (2012), Typology approach for building stock energy assessment. Final project report, Intelligent Energy Europe, 14-15.
[12] IEE (2010), Summary of the recommendations of the ASIEPI project, Intelligent Energy Europe
[13] IEE (2009), A set of reference buildings for energy performance calculation studies, Intelligent Energy Europe
[14] Kurnitskia, J., Saari, A., Kalamees, T., Vuolle, M., Niemela, J., Tark, T. (2011), Cost optimal and nearly zero (nZEB) energy performance calculations for residential buildings with REHVA definition for nZEB national implementation, Energy and Buildings, 43(11), 3279-3288
[15] Olofsson, T. & Mahlia, T.M.I. (2012), Modeling and simulation of the energy use in an occupied residential building in cold climate, Applied Energy, 91(1), 432-438.
[16] Stegmann, M.B. & Gomez, D.D. (2002), A Brief Introduction to Statistical Shape Analysis, Lecture note, Informatics and Mathematical Modelling, Technical University of Denmark, DTU
[17] Tommerup, H. & Svendsen, S. (2006), Energy savings in danish residential building stock, Energy and Buildings, 38(6), 618-626.
[18] Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor (2001), Active Appearance Models, IEEE transactions on pattern analysis and machine intelligence, 23(6), 681-685