International Journal of Advanced Culture Technology Vol.12 No.3 434-440 (2024)
DOI https://doi.org/10.17703/IJACT.2024.12.3.434

IJACT 24-9-52

Analysis of Female Lower Body Shapes for the Development of Slacks Patterns: Exploring Body Clusters Using Machine Learning

[1]Ji Min Kim

[1] *Associate Professor, Dept. of Fashion Design, Gangneung-Wonju National Univ., Korea*
kimjm@gwnu.ac.kr

Abstract

SIZE KOREA updates body measurement data every five years, providing essential information for the fashion industry. This anthropometric data is widely used to diagnose consumer body shapes and develop optimal clothing sizes. Artificial intelligence, particularly machine learning, excels in predicting such body shape classifications. This study seeks to enhance the suitability of clothing design by applying the new analytical methodology of machine learning techniques to better capture and classify the unique body shapes of Korean women.

In this study, machine learning techniques such as K-means clustering, Silhouette analysis, and Decision Tree analysis were used to classify the lower body shapes of Korean women in their twenties and identify standard body shapes useful for slacks design. The results showed that the lower body of the age group could be classified into three categories: 'small stature' (the majority), 'tall with an average lower body volume,' and 'medium height with a fuller lower body' (the smallest share). The three-cluster approach is validated through Silhouette analysis, which minimizes misclassification.

Decision Tree analysis then further defines the criteria for these clusters, highlighting waist height and hip depth as the most significant factors, achieving a classification accuracy of 90.6%. While this study is not directly related to Robotic Process Automation, its detailed analysis of body shapes for slacks patterns can aid RPA in clothing production. Future research should continue integrating machine learning in human body and fashion design studies.

Keywords: *Lower Body Shapes, K-means, RPA, Slacks Patterns, Body Clusters*

1. INTRODUCTION

SIZE KOREA, a government agency, updates body measurement information every five years. This data serves as a foundational element for the fashion industry. Since 1979, data on more than120,000 individuals have been disclosed over eight investigations up to 2023. This study analyzes data from the most recent, the eighth investigation, specifically focusing on 575 women in their twenties. The study is to identify standard body shapes useful for the design and production of slacks for women in this age group.

SIZE KOREA categorizes the physiques of Korean women into four shapes, considering the full

Manuscript received: July 11, 2024 / revised: August 14, 2024 / accepted: September 5, 2024
Corresponding Author: kimjm@gwnu.ac.kr
Tel:+82-33-640-1656, Fax: +82-33-640-1691
Associate Professor, Dept. of Fashion Design, Gangneung-Wonju National Univ., Korea

body—both the upper and lower halves. Previous studies, such as those by Kim *et al.* (2022), analyzed the lower body into four shapes, using factor analysis[1]. However, recent advancements in artificial intelligence (AI) and machine learning (ML) have led to new findings that differ from those of earlier research.

Consequently, this research employs various machine learning techniques, including K-means, Silhouette analysis, and Decision Tree, to analyze the latest body measurement investigation data. The goal is to classify and present body shapes that can help develop the most suitable slacks patterns for women in their twenties. Generally, artificial intelligence is composed of two elements: Machine Learning and Robotic Process Automation (RPA). Machine Learning analyzes data to classify shapes and instructs robots on their tasks. In contrast, RPA encompasses a wide range of applications, not only in manufacturing machinery but also in office automation tools and software or applications. Identifying suitable slacks patterns falls under the domain of ML, while using these patterns to produce actual garments pertains to the RPA process. Categorizing the body shapes of Korean women is a task for ML, and using this classification to create slacks is within the domain of RPA. ML and RPA are so intrinsically linked that the generation of accurate pattern classifications by ML is essential for RPA to efficiently advance the fashion industry. If this machine learning study accurately identifies body clusters, then robots in clothing manufacturing companies will produce better-fitting products, which make consumers and fashion companies generate greater satisfaction.

2. LITERATURE REVIEW

Numerical data on human size and body shape serve as fundamental information, widely utilized in numerous studies. However, a review of the existing literature reveals two predominant research trends. The first trend investigates changes in human body measurements over time, while the second seeks to analyze these measurements to identify consistent patterns or shapes. Relating to this study, previous research on the latter is summarized as follows.

2.1 Body Shape Typologies

Studies related to body shape typologies attempt classification in relation to specific purposes of clothing. Specifically targeting women's clothing, researchers classify female body shapes in various ways according to their research objectives, as shown in Table 1. It presents a very interesting finding. Many studies, each with their specific objectives, classify the body shapes of Korean women into three or four distinct shapes. Most of them adopt Factor Analysis for their classification[2-6]. Among the many existing studies reviewed, the one most closely related to this research is by Kim *et al.* They found four body shapes for designing slacks for young women: petite lower body with short legs; long lower body with medium built; obese abdomen and legs; short and slim lower body. The study employed the classic factor analysis approach, too.

Table 1. Summary of Previous Research Classifying Body Shapes

Author(s)	*Pub. Year*	*Purpose of Study*	*Body Parts Analyzed*	*Results: Classified Body Shapes*
Yoon, Hong	2023	Classifying obese female body for avatar modeling	Whole body	(4) Curved; well-developed lower body/stooped; upright/flat; twisted/Well-developed upper body, long legs
Cha	2024	Tall women's body shapes	Whole body	(3) Thin & short body shape; tall average body shape; thick & short obese body shape
Kim	2022	Developing clothing patterns	Upper body	(4) Ave. body/torso; large body/obese; straight body/short; tall/curvy
Rui et al.	2023	Classification and analysis of torso shapes	Upper body	(3) Long upper body, flat back/hips; most curved back/forwarded upper body, least protruding abds./most sloped shldrs.; short upper body, most protruding hips/largest body
Koo	2023	Developing	Lower	(4) Average in length/size, most curved,

		better-fitting uniform skirt	body	longest W-H; thin/longest in length, shortest W-H; short/largest in size, least curvy, short W-H; shortest/thinnest, long W-H
Kim et al.	2022	Slacks pattern development	Lower body	(4) Petite lower body with short legs; long lower body with medium built; obese abdomen and legs; short and slim lower body

2.2 Lessons Learned from Literature Review

The conventional method productively examined appropriate body shapes for Korean women. However, while factor analysis provides the factor loading values for individuals, it does not explicitly indicate which cluster each individual belongs to. It is certainly a limitation of this method. In contrast, the K-means technique not only categorizes body shapes into several clusters but also provides the threshold values for the variables that define these divisions. Consequently, this approach facilitates the easy prediction of the body shape group to which each instance belongs. The distinction between this study and previous research lies in the choice of analytical methodologies. Specifically, this research seeks to identify body shapes of the young women more suitable for developing clothing. Learning from the previous literature, this research employs new analytical methodologies to better capture and classify the unique body shapes of the age group, aiming to improve the suitability of clothing patterns.

3. METHODOLOGY

Extensive anthropometric data is currently utilized for diagnosing consumer body shapes and developing optimal clothing sizes. Predicting these body shape classifications is a field where artificial intelligence, particularly machine learning, excels. This study aims to apply machine learning techniques, with the ultimate goal of aiding the development of better-fitting slacks patterns for women in their 20s.

Both supervised learning and unsupervised learning approaches are used for classification. The primary distinction between supervised and unsupervised learning is the presence of a target variable or label, which functions as the independent variable in traditional statistics. If a target variable (label) is present, supervised learning techniques are applied; otherwise, unsupervised learning techniques are utilized.

In this study, since there are no specifically designated or certified body shapes for slacks patterns, an unsupervised learning approach is adopted. Considering alternative methodologies like t-SNE, Louvain Clustering, PCA, and MDS, this study employs the K-means method due to its status as the most widely used and universal unsupervised classification technique.

The K-means algorithm initially creates several arbitrary clusters and sets centroids for each cluster. It then calculates the distance of each data point from these centroids. By continuously adjusting the centroids, the algorithm minimizes the total distance between the data points and their respective centroids, ultimately selecting the optimal clusters. This trial-and-error approach to finding optimized clusters is a typical method used in machine learning.

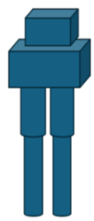
3.1 Data and Recent Changes in Korean Women's Lower Body Measurements

The key lower body measurement items from the 5th, 7th, and 8th Investigations were selected, totaling eighteen items. These include measurements for height (stature, waist, hip, crotch, knee), breadth (waist, hip), depth (waist, hip), circumference (waist, hip, thigh, knee, calf, minimum leg, maximum ankle), and length (outside leg, crotch). The measurements of these key items were compared and analyzed across different investigation cycles, and the items with significant changes are presented in Table 2.

The visualization of these data corresponding to specific body features is illustrated in the left columns of Table 2. In this figure, the left side represents the human form, while the right side visualizes the data used in the body measurement investigation. For example, the variables for waist height, waist breadth, and waist depth form a rectangular prism representing the waist area, and similarly, the hip height, hip width, and hip depth form another rectangular prism. Additionally, the crotch height and thigh circumference create two cylindrical shapes, and the knee height and maximum ankle circumference are depicted by the remaining two cylinders.

Table 2. Visualization of Data for Lower Body Compared with Human Body Shape

unit: cm

Human Body	***Visualization of Data***	**Measured Items**	**Mean: 8th Meas.**	***Difference: 8th-5th***	**Difference: *8th-7th***
		Stature	161.25	+1.21***	+0.37
		Waist height	98.31	-1.30***	+1.19***
		Waist breadth	25.53	+1.50***	+0.12
		Waist depth	18.26	+0.50***	+0.68***
		Hip height	78.28	+0.32***	+0.91***
		Hip width	33.49	+1.07***	+0.89***
		Hip depth	23.51	+2.56***	+2.26***
		Crotch height	73.13	+0.86***	-0.66
		Thigh circumference	56.12	+2.26***	+1.31***
		Max. ankle circumference	23.27	-0.08	+0.15***

Note: Asterisks(***) denote the measures are significantly different at the level of 95% or higher.
Human body generated by CLO 3D using the 8th anthropometric data
No. of participants: 5th (N=692), 7th (N=668), 8th (N=575)

3.2 Machine Learning Analysis

Subsequently, machine learning techniques, specifically K-means clustering, were employed to categorize the lower body shapes of women in their twenties. Silhouette analysis was conducted to validate the three-cluster approach and minimize misclassification. Additionally, decision tree analysis was utilized to clearly define the criteria for each cluster, identify the variables most influential in body shape classification, and demonstrate that the classification accuracy achieved by the decision tree analysis is at a very high level.

4. RESULTS AND DISCUSSION

This study compares the 8th round measurement (2020-2023) of SIZE KOREA with the 5th round measurement (2003-2004) and the 7th round measurement (2015). The measurements of the lower body of women taken in the 5th, 7th, and 8th cycles reflect a time gap of approximately 20 years and 5 years, respectively, when viewed from the 8th cycle. The most significant change in body shape over these 20 years is observed in hip depth. In 2003-2004, the average hip depth was 20.95 cm, which increased to 21.25 cm in 2015, and further increased to 23.51 cm in 2020 and 2023. Therefore, the trend of increasing hip depth is a crucial consideration for slacks patterns. The next most significant change is in thigh circumference. Twenty years ago, during the 5th cycle measurement, it was 53.86 cm, which increased by 2.26 cm to 54.81 cm in the 7th round, and further increased by 1.31 cm in the 8th round, reaching an average of 56.12 cm.
On the other hand, the average values related to height, such as waist height, hip height, crotch height, and knee height, have increased compared to 20 years ago. However, compared to five years ago, these height-related measurements have not shown significant growth and, in some cases, have even decreased. Therefore, it can be said that the body shape of Korean women in their 20s is changing more towards a wider, thicker, and rounder shape rather than a change in vertical length.

4.1 Initial Stage of K-means Analysis

In the dataset of 20 to 29 years old female from the 8th Anthropometric Investigation, the variables listed in Table 2 are referred to as features in machine learning terminology. The K-means clustering was performed using these features. The analysis was conducted using the Brightics AI, a machine learning tool developed by Samsung SDS. As a result, K-means classifies the lower body shapes into two categories, as shown in Table 3.

The results indicate that Cluster 2, compared to Cluster 1, has only a slightly lower knee height but exhibits larger values in all other measured dimensions. Consequently, Cluster 1 is classified as the generally smaller body shape, while Cluster 2 is categorized as the larger shape. Out of the total 575 participants analyzed, 272 individuals, accounting for 47.3%, are classified under the "smaller body shape," and 303 individuals, representing 52.7%, are classified as the "larger body shape."

4.2 Silhouette Analysis

The accuracy of this classification is also examined by the silhouette scores. The silhouette score is a measure of the quality of clustering; a score close to 1 indicates that the data is well-matched to its cluster, a score of 0 suggests that the data is on the edge of its cluster, and a negative score indicates improper classification within the cluster. The left two columns of Table 3 represents silhouettes as pictograms. The instances that are classified incorrectly are marked to the left of the centerline in the tail part of the silhouette.

In the K-means clustering technique, researchers can arbitrarily assign the number of clusters to be classified. As demonstrated in Table 3, the number of clusters are set at two, three, four, and five, respectively. The silhouette values are then used to identify the extent of misclassification within each grouping. It is observed that misclassification is minimized when the number of clusters is three. To this end, the number of clusters is set to three, and a secondary K-means cluster analysis is conducted.

Table 3. Number of Misclassification

Silhouette for Small Body Shape	*Silhouette For Larger Body Shape*	**No. of Misclassified for 2 Clusters**	**No. of Misclassified for 3 Clusters**	**No. of Misclassified for 4 Clusters**	**No. of Misclassified for 5 Clusters**
		3 5	***3 3***	4 6	4 3

4.3 Secondary K-means Analysis

As Silhouette analysis indicates, the secondary classification is conducted by setting the number of clusters to three at the K-means analysis. Table 4 presents the average and standard deviation of measurement items for the lower body of women in their twenties, segmented into three clusters. The highest values across the measurement items for each cluster are highlighted in bold. As a result, Cluster 1 exhibited the largest average values in length-related items, while Cluster 2 had the largest average values in items related to the volume of the lower body. Cluster 3, conversely, showed the smallest lower body sizes across all items. Therefore, Cluster 1 can be characterized as 'tall stature with an average lower body,' Cluster 2 as 'medium height with a fuller lower body,' and Cluster 3 can simply be described as 'small stature,' with this shape comprising the majority of the young female population studied.

Table 4. Characteristics of Body Shape Categories

unit: cm

Measurement Item	*Cluster 1*	*Cluster 2*	*Cluster 3*
Waist height	***101.98*** (2.61)	97.58 (3.79)	95.93 (2.59)

Waist breadth	25.40 (1.73)	***29.02*** (1.93)	24.19 (1.62)
Waist depth	17.77 (1.67)	***22.11*** (2.28)	17.07 (1.67)
Hip height	***81.45*** (2.59)	77.59 (3.23)	76.25 (2.41)
Hip width	33.84 (1.48)	***35.87*** (1.86)	32.27 (1.31)
Hip depth	23.17 (1.63)	***27.27*** (2.16)	22.21 (1.57)
Crotch height	***76.42*** (2.45)	72.86 (3.23)	71.25 (2.39)
Thigh circumference	55.99 (3.56)	***63.43*** (3.99)	53.19 (3.37)
Knee height	***43.17*** (1.51)	40.99 (1.72)	40.24 (1.58)
Knee circumference	35.66 (1.75)	***38.49*** (2.10)	33.87 (1.59)
Max. ankle circumference	23.66 (0.91)	***24.29*** (1.02)	22.55 (0.81)
No. of Instances	196 (34.1%)	111 (19.3%)	268 (46.6%)

Note: Numbers in the parenthesis are standard deviations associated with the mean.

4.4 Decision Tree Analysis

The next step is to explore the criteria used to create the clusters. To achieve this purpose, decision tree analysis is employed. Decision trees are easy to understand and interpret. They provide a clear visualization of the cluster-making process. They often provide insights into which variables are most important in classification.

A total of 575 participants are initially divided into groups based on whether their waist height is less than 99.35 cm. Subsequently, individuals within the group with a smaller waist height are further divided based on whether their hip depth is less than 25.2 cm. Participants with a hip depth less than 25.2 cm are classified into Cluster 3, characterized as 'small stature.' Those with a hip depth greater than 25.2 cm are then classified into Cluster 2, described as 'medium height with a fuller lower body.'

On the other hand, 222 participants whose waist height exceeds 99.3 cm are subdivided into two groups. Participants with a waist depth less than 20.6 cm are classified into Cluster 1, characterized as 'tall with an average lower body volume.' Those not meeting this criterion are classified into Cluster 2, previously described as 'fuller-bodied.'

In this classification, the most influential feature is waist height, which accounts for 48% of the relative impact. The next most significant feature is hip depth, with a relative influence of 30%. The relative influence of waist depth is 13%. Other features had only a marginal impact. The classification accuracy or CA of this decision tree analysis is 90.6%, indicating a high level of precision. F1-score of 92.6% is high enough, too.

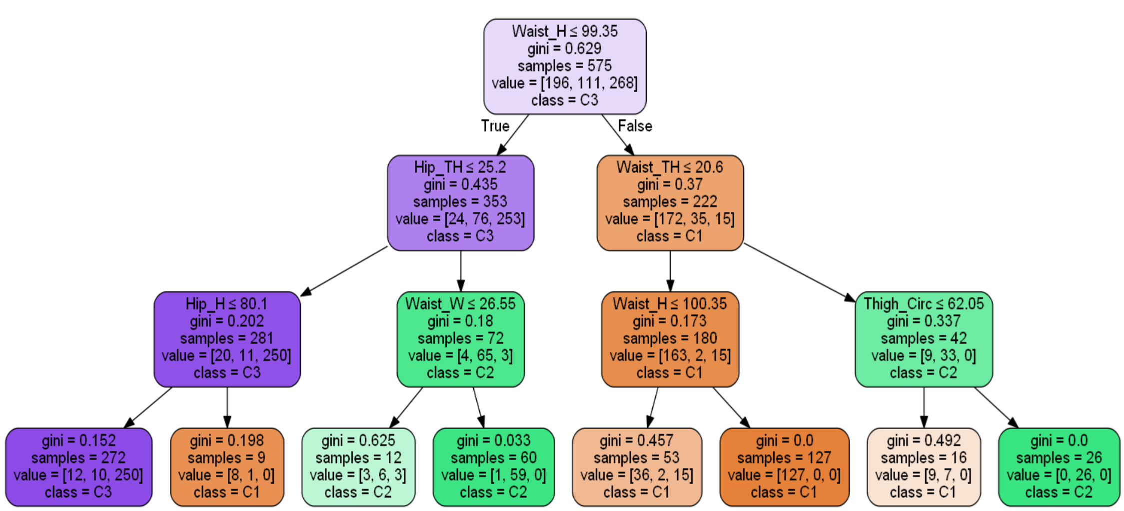


Figure 1. Decision Tree for Classifying Lower Body Shapes

5. CONCLUSION

Like previous researches, this study analyzed anthropometric data of Koreans measured and published by the government agencies. It classified the body shapes of South Korean women in their twenties into three categories, akin to other studies that categorizes Korean women's body shapes into three or four shapes. However, unlike other studies, this research employed novel machine learning analysis techniques.

The K-means clustering method categorized the lower bodies of women in their twenties into three shapes: the 'small stature' shape constitutes the majority, followed by the 'tall with an average lower body volume' shape. The 'short with a fuller lower body' shape has the smallest share.

Why are three shapes more appropriate than two, four, or five shapes? This was a question difficult to answer in previous studies. However, this study demonstrated, by analyzing the Silhouette score, that three clusters minimized the number of misclassification instances.

Additionally, by incorporating decision tree analysis, this study quantitatively and clearly defined the criteria for categorizing these three clusters. waist height and hip depth play the most significant roles in classifying the lower body shapes of women in their twenties in Korea. The classification accuracy for these categorized body shapes was high at 90.6%.

Although this study is not directly related to artificial intelligence or Robotic Process Automation (RPA), the thorough analysis of lower body shapes for better-fitting slacks patterns can significantly contribute to the advancement of RPA in garment production. From this perspective, future research should actively and continuously incorporate new technologies such as machine learning into the study of human body shapes and fashion design.

REFERENCES

[1] J.M. Kim, S.H. Um, Y.S. Lee, Y.M. Kim, and H.R. Woo, "Analysis of the Changes in Lower Body Measurements and Shapes of Women in Their 20s for Slacks Pattern Development", *Journal of Fashion Business*, Vol. 26, No. 1, pp. 30-40, Feb. 2022. https://doi.org/10.12940/jfb.2022.26.1.30

[2] J.W. Yoon, E.H. Hong, "A Study on the Body Type for Avatar Modeling of Middle-Age Obese Women", Culture and Convergence, Vol. 45, No. 1, pp. 767-780, Oct. 2023. https://doi.org/10.33645/cnc.2023.10.45.10.767

[3] S.J. Cha, "Body Types of Tall Women Over 165cm", *Journal of the Korean Society of Clothing and Textiles*, Vol. 48, No. 3, pp. 528-542, Jun. 2024. https://doi.org/10.5850/JKSCT.2024.48.3.528

[4] J.H. Kim, "An Analysis of the Torso Shapes of Elementary School Girls", *Bulletin of Korean Society of Basic Design & Art,* Vol.23, No.1, pp. 121-132, Feb. 2022. http://dx-doi-org.libproxy.gwnu.ac.kr/10.47294/KSBDA.23.1.9

[5] Y. Rui, E.J. Ryu, and H.K. Song, "Classification of Torso Shapes of Women Aged 35-54 -Based on Measurements Extracted from the 8th Size Korea Scans", *The Korean Fashion and Textile Research Journal,* Vol. 25, No. 5, pp. 603-614, Oct. 2023. https://doi.org/10.5805/SFTI.2023.25.5.603.

[6] M.R. Koo. "A Study on the Lower Body Type Classification of Female High School Students for School Uniform Skirts Design", *The Journal of Humanities and Social Sciences 21,* Vol. 14, No. 3, pp. 2479-2491, Jun. 2023. http://dx.doi.org/10.22143/HSS21.14.3.174