

**PENGELOMPOKKAN CITRA AKSARA BALI MENGGUNAKAN SINGLE
LINKAGE HIERARCHICAL CLUSTERING BERDASARKAN HASIL EKSTRAKSI
CIRI FREEMAN CHAIN CODE**

Oleh

David Thanlian Kurniawan

(21/475808/PPA/06139)

Daun lontar merupakan bentuk media yang digunakan pada orang zaman dahulu untuk berkomunikasi secara tulisan, sehingga tidak menutup kemungkinan terdapat suatu informasi yang terkandung pada daun lontar tersebut menjadi menghilang. Dalam penelitian ini, telah didapatkan set data segmentasi aksara Bali dalam bentuk citra digital, dengan setiap unsur karakter masih perlu diproses lebih lanjut.

Clustering menjadi salah satu cara yang digunakan untuk menyelesaikan permasalahan terhadap jumlah jenis aksara dari set data, sehingga aksara daun lontar tersebut tidak hanya disimpan sebagai bentuk citra digital saja, tetapi setiap aksara yang terkandung dalam segmentasi daun lontar dapat berkontribusi sebagai model database training.

Sebelum dilakukan nya *clustering*, mula-mula set data dari citra aksaraa Bali akan melalui proses *preprocessing* yang melibatkan proses segmentasi dan *noise removal*. Selanjutnya memasuki proses ekstraksi ciri yang berbeda, yakni Freeman Chain Code yang memiliki kehandalan dalam melakukan retrieving shape, Count of Foreground yang merupakan bentuk ekstraksi ciri berbasis *region*, dan gabungan antara Freeman Chain Code dan Count of Foreground. Kemudian dilakukannya pengelompokan dengan Agglomerative Hierarchical Clustering Single-Linkage untuk masing-masing bentuk ciri. Setelah itu dilakukannya pengujian menggunakan *silhouette coefficient* untuk melihat kombinasi ekstraksi ciri, fungsi *similarity*, serta ukuran window ciri yang menghasilkan kualitas kelompok yang baik.

Melalui percobaan yang telah dilakukan, *boundary base* yang digunakan sebagai ekstraksi ciri bagi citra aksara Bali lebih cocok dibandingkan dengan *region base*. Tetapi, hasil pengelompokan dengan menggunakan *single-linkage hierarchical clustering* masih belum bisa memberikan hasil yang optimal

Kata Kunci: Clustering, Single Linkage, Hierarchical Clustering, Ekstraksi Ciri, Freeman Chain Code. Count of Foreground

BALINESE SCRIPT IMAGE GROUPING USING SINGLE LINKAGE HIERARCHICAL CLUSTERING WITH FREEMAN CHAIN CODE FEATURE EXTRACTION

Oleh

David Thanlian Kurniawan

(21/475808/PPA/06139)

Lontar leaves are a form of media used by ancient people to communicate in writing, so it is possible that the information contained in the palm leaves will disappear. In this study, the Balinese script segmentation data set has been obtained in the form of digital images, with each character element still needing further processing.

Clustering is one of the ways used to solve the problem of the number of types of characters from the data set, so that the lontar leaf script is not only stored as a form of digital image, but each character contained in the lontar leaf segmentation can contribute as a training database model.

Prior to clustering, the data set from the Balinese script image will first go through a preprocessing process which involves segmentation and noise removal process. Then it enters a different feature extraction process, Freeman Chain Code which has the reliability in retrieving shapes, Count of Foreground which is a form of region-based feature extraction, and a combination of Freeman Chain Code and Count of Foreground. Then do the grouping with Agglomerative Hierarchical Clustering Single-Linkage for each feature type. After that, testing was carried out using the silhouette coefficient to see the combination of feature extraction, similarity function, and feature window size to produced good group quality.

Through the experiments that have been carried out, the boundary base used as feature extraction for Balinese script images is more suitable than the region base. However, the results of grouping using single-linkage hierarchical clustering still cannot provide optimal results

Keyword : Clustering, Single Linkage, Hierarchical Clustering, Feature Extraction, Freeman Chain Code. Count of Foreground.