ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Structural Substitution and Stylistic Integration in Film and Television Adaptations of Chinese and Foreign Classical Literature

Li Fan[1,*] *and Yongfei Ye*[2]

[1] School of Literature, Hebei North University, Zhangjiakou, Hebei, 075000, China
[2] School of Information Science and Engineering, Hebei North University, Zhangjiakou, Hebei, 075000, China

SUMMARY: *This paper introduces computational narrative analysis and visual recognition algorithms to combine quantitative and qualitative research on film and television adaptations of Chinese and foreign classic literature. The film and television works of literary adaptation of Pride and Prejudice and Dream of the Red Chamber are selected as examples, and through TF-IDF, LDA thematic model and deep learning methods, the plot structure, emotional arc and thematic model of the script and the original are compared to each other, to reveal the law of narrative structure replacement in the adaptation; at the same time, the combination of image features combining the spatial pyramid model and contextually relevant histogram are utilized to construct the work At the same time, the combination of spatial pyramid model and context-dependent histogram image features is used to construct a work style analysis model, and the key frames of the film are identified with typological visual elements. From the results of the study, it can be seen that there are seven themes in Chinese and foreign literary works, the warm tone of film and television works accounted for more than 20%, the cold tone occupies a small portion, the film and television content is dominated by the warm tone, in the emotional fragments of panic and excitement, the sound effects and the background music production are mostly using percussion musical instruments, in the mood of relaxation and sadness, the film and television works of the orchestral musical instruments are more used, the tempo of the camera transition are relatively smooth.*

KEYWORDS: *structural substitution; TF-IDF; LDA theme model; visual element recognition; film and television works*

1 Introduction

The wave of globalization promotes the deep interaction of literature and culture [1]. As a medium integrating visual art, narrative skills and cultural elements, film and television works have become an important carrier of cultural dissemination [2, 3]. Classic literature adapted into film and television works can not only provide audiences with a new visual experience, but also reproduce and reconstruct the profound connotation of literary classics [4, 5]. Film and television works skillfully integrate the classic flavor of the original literature with modern cultural elements to create new values and meanings, and reflect the unique advantages of film and television works as a tool for cultural exchange [6, 7]. Exploring the adaptation of literary classics in film and television not only involves the cross-border integration of literature and video media, but also deeply reflects the mutual influence and

*13931319190@163.com
https://doi.org/10.65102/is2026426
Published: 30 April, 2026

development of two different cultures in the era of globalization and informationization [8, 9]. The creation of film and television works adapted from literary classics can reveal the cultural qualities, aesthetic differences, and social effects of classic literary responses, aiming to enhance the depth and breadth of cultural communication and promote the symbiosis and integration of cultural diversity [10, 11].

From literature to movie and television works, from readable text to visual text, there are confrontation and integration, and existing studies mostly compare literature and movie and television works, focusing on their differences and neglecting the intertextual phenomenon between them [12, 13]. In the media era with the film and television adaptation of classic literature as the entry point, it is essential to study the structural and stylistic relationship between literature and images, which is not only a problem of media conversion, but also a problem of literary development. This not only allows us to better understand the significance of literary pictorialization, but also provides better ideas for structural replacement and stylistic integration and development.

In this paper, we take the literary adaptation film and television works Pride and Prejudice and Dream of the Red Chamber to construct the multimodal dataset of Chinese and foreign classic literary adaptation film and television works, and with the help of natural language processing technology and computer vision method, we reveal the structural adjustment and stylistic adaptation features in the process of adaptation of Chinese and foreign classic literary works. First, the TF-IDF method is used to represent and preprocess the text of Chinese and foreign classic literary works, then the LDA model is used to model the themes of different literary and film adaptations, and the deep learning model is used to compute the emotional arcs of literary and adapted film and television works, so as to explore the structural narrative replacement of adapted film and television works. Secondly, computer vision technology is used to extract the texture and color elements of Chinese and foreign classic literature adapted film and television works, and based on the BPM method to extract the rhythmic characteristics of film and television audio, to explore the stylistic integration of adapted film and television works.

2 Measurement of structural replacement and stylistic integration in literary adaptations for film and television

2.1 Construction of the dataset of Chinese and foreign classic literary adaptations of film and television works

2.1.1 Textual data on Chinese and foreign classic literature

The study takes the foreign movie Pride and Prejudice and the domestic movie and television works of Dream of the Red Chamber as examples, and collects the script texts of Pride and Prejudice and Dream of the Red Chamber movie and television works, and through the subtitle files, official scripts, the length of the footage, as well as the full text of the corresponding original novels, it collects a total of 547 points of audio files and 267 text files after the collation.

2.1.2 Video data of adapted film and television works

From Pride and Prejudice, Dream of the Red Chamber film and television works, the film and television drama will be extracted key frames according to the 40-minute time interval or based on the scene segmentation, to get the expression of emotions in film and television works, textual feelings and the artistic embodiment of the literary adaptation of film and television.

2.2 NLP-based model for analyzing the narrative flow of works

2.2.1 Representation of literary works

TF-IDF is a common text representation in information retrieval that utilizes the importance of words compared to One-Hot. The word frequency (TF) is the frequency of each word in the current document, and the formula is shown in (1). In order to avoid counting the deactivated words that have no real meaning, these deactivated words should be eliminated when doing preprocessing in general, so it is necessary to assign the corresponding weight to each word, i.e., Inverse Text Frequency Index (IDF), the formula is shown in (2). In the denominator, 1 is added to avoid the situation that the word does not exist in the document.IDF utilizes the deactivated words will be in all the corpus in a high frequency of the characteristics, so a word in all the documents, the more common, the lower its IDF, the more unimportant these words are, that is, the lower the weight is given.The calculation method of the TF-IDF is to multiply the frequency of the word and its corresponding inverse text frequency index, the formula is shown in (3). The formula is shown in (3).

Word Frequency:

$$\text{Word Frequency(TF)} = \frac{\text{The number of times a word appears in the current document}}{\text{Total number of words in the current document}} \tag{1}$$

Inverse text frequency index:

$$\begin{aligned}&\text{Inverse text frequency index(IDF)}\\&= \log\left(\frac{\text{Total number of documents in the corpus}}{\text{Total number of documents in which the word appears } +1}\right)\end{aligned} \tag{2}$$

$$\text{TF} - \text{IDF} = \text{Word Frequency(TF)} \times \text{Inverse text frequency index(IDF)} \tag{3}$$

The advantage of the TF-IDF algorithm is that it is relatively simple to implement and easy to understand. The disadvantage is that it is very dependent on the corpus, and it is not accurate enough to determine the importance of words by calculating word frequency. In addition, in this algorithm, the different positions of the words have no effect on the results, and the correlation between the contextual words in the text is not taken into account.

2.2.2 LDA-based theme modeling of works

The implicit Dirichlet distribution, also known as LDA, is a widely used document topic generation model. The idea of this model is to utilize the fact that topics contain the core content of a document, and a document can have one or more topics.The assumption of LDA is that each document can be regarded as a probability distribution in the set of topics, which obeys the Dirichlet distribution, and each topic can be regarded as a probability distribution in the set of words, which also obeys the Dirichlet distribution. If $(\alpha_1, \alpha_2, \ldots, \alpha_d)^T$ obeys the parameter $(\theta_1, \theta_2, \ldots, \theta_d)^T$ of the Dirichlet distribution, then its probability distribution expression is shown in (4):

$$p\left(\theta_1,\theta_2,\ldots,\theta_d \mid \alpha_1,\alpha_2,\ldots,\alpha_d\right)=\frac{\Gamma\left(\sum_{k=1}^{d}\alpha_k\right)}{\sum_{k=1}^{d}\Gamma\left(\alpha_k\right)}\prod_{k=1}^{d}\theta_k^{\alpha_k-1} \tag{4}$$

From the idea of the model, we can obtain Eq:

$$p(\text{Word} \mid \text{Document})=\sum_{\text{Topic}} p(\text{Word} \mid \text{Theme})\times p(\text{Topics} \mid \text{Documents}) \tag{5}$$

The LDA model is shown in Figure 1, where D denotes the number of documents in the document set, N represents the number of words, and K denotes the number of topics contained in the documents, which need to be specified first. α and β are a priori parameters given by past experience, θ_d denotes the distribution of topics in the d th document, and φ_k denotes the distribution of words in the k th topic. From the assumptions of LDA, it follows that obeys $Dirichlet(\alpha)$ and obeys $Dirichlet(\beta)$. $z_{d,n}$ denotes the topic of the n th word in the d th document, and is able to obtain the distribution of the topic of the n th word in the d th document from the distribution of the topic of the d th document, obeying $Multionmial\left(\theta_d\right)$. $w_{d,n}$ represents the n th word of the d th document, which obeys $Multionmial\left(\varphi_{z_{d,n}}\right)$ in that topic.

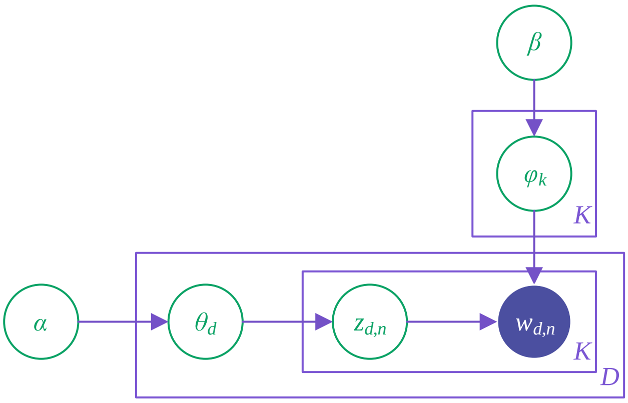


Figure 1: LDA model

The generation steps of LDA are: firstly, obtain the number of words in the d st document N; then, get the parameters of topic-document distribution and topic-word distribution according to $\theta_d \sim Dirichlet(\alpha)$ and $\varphi_k \sim Dirichlet(\beta)$, respectively; then get the topic of the word according to $z_{d,n} \sim Multionmial\left(\theta_d\right)$; and finally, get the word according to $w_{d,n} \sim Multionmial\left(\varphi_{z_{d,n}}\right)$. Therefore, the joint probability formula of LDA model can be obtained:

$$p\left(w_d,z_d,\theta_d,\varphi \mid \alpha,\beta\right)=\prod_{n=1}^{N} p\left(w_{d,n} \mid \varphi_{z_{d,n}}\right)p\left(z_{d,n} \mid \theta_d\right)p\left(\theta_d \mid \alpha\right)p\left(\varphi_{z_{d,n}} \mid \beta\right) \tag{6}$$

The model usually uses Gibbs sampling to do parameter estimation, and the theme of each word can be obtained after inferring the parameters θ and φ.

2.2.3 Calculation of the emotional arc of a literary work

1) Computational steps

The specific steps of deep learning-based sentiment analysis are shown in Figure 2. Firstly, the text is preprocessed, and then the words are textually represented, and the methods include Word2vec, BERT, etc. Then the features are further extracted using neural network algorithms, and finally, the model is accessed to the full connectivity layer and softmax to output the probability of each classification, so that the sentiment classification can be obtained.

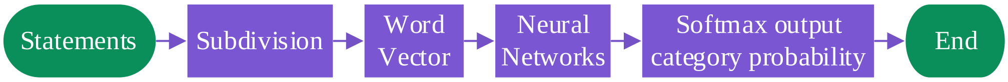


Figure 2: Flowchart of the sentiment analysis method based on deep learning

2) Computational methods

(1) Recurrent Neural Networks

Recurrent neural network RNN is mainly used in natural language processing.In most of the neural network algorithms, the data is inputted, passed through the hidden layer, and finally the result is outputted, but the process may have some temporal order related to the

Therefore, RNN has the following formula:

$$s_t = \tanh\left(x_t \times U + s_{t-1} \times W + b\right) \tag{7}$$

where, denotes the word vector of the input text at the moment of t after disambiguation, s_t denotes the hidden state at the moment of t, W and U stand for both represent the weight matrices, b denotes the bias term, and tanh denotes the activation function tanh.

The reason that RNN can be widely used in the field of natural language processing is because each word in the sentences people speak everyday is in order, and when we predict the result, we need to take the previous features into account, and the prediction will be more accurate.

(2) Long and short-term memory networks

The cell structure of LSTM is shown in Fig. 3. Among them, C denotes the memory state, which records the historical information of each moment and is the long-time memory. s_t denotes the hidden state at the moment of t, which is short-time memory. First, the input x_t and the hidden state s_{t-1} of the previous moment are jointly used as inputs to compute the probability vector i_t, which is known as the input gate, and determines how much information from the inputs can be transmitted to the unit state, with the following formula:

$$i_t = \sigma\left(W_{ix} x_t + W_{is} \times s_{t-1} + b_i\right) \tag{8}$$

where σ denotes the sigmoid function, W_{ix} represents the weight matrix corresponding to the input x at the moment t, and b_i is the bias term corresponding to the calculation of i_t.

Then, a vector of candidate terms $\tilde{C}_t$ is constructed for the subsequent computation of the cell state with the following formula:

$$\tilde{C}_t = \tanh\left(W_{cx}x_t + W_{cs}s_{t-1} + b_c\right) \tag{9}$$

The forgetting gate is then constructed to decide what information to discard as well as what to keep, with the following formula:

$$f_t = \sigma\left(W_{fx}x_t + W_{fs}s_{t-1} + b_f\right) \tag{10}$$

where f_t denotes the forgetting gate factor.

The cell state is updated to calculate how much of the previous moment's memory state C_{t-1} can be retained in the current moment's memory state C_t, with the following formula:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{11}$$

Finally, the final state is computed:

$$o_t = \sigma\left(W_{ox}x_t + W_{os}s_{t-1} + b_o\right) \tag{12}$$

$$s_t = o_t \tanh\left(C_t\right) \tag{13}$$

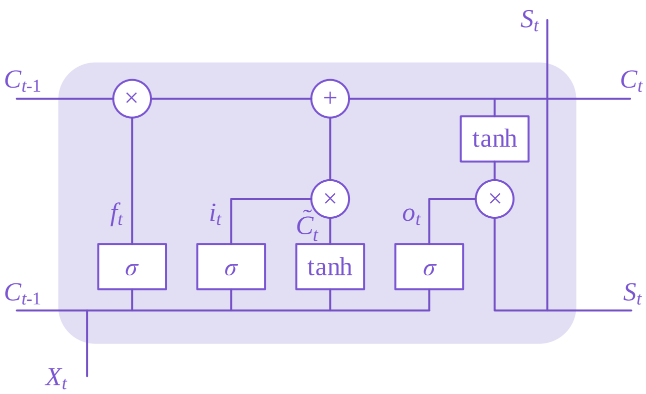


Figure 3: Unit structure diagram of LSTM algorithm

For text classification based on LSTM, a layer of softmax is added at the end of the LSTM model and the final output is:

$$y_t = \text{softmax}\left(W_m s_t + b_m\right) \tag{14}$$

where W_m is the corresponding weight vector and b_m is the bias term.

2.3 A computer vision-based model for analyzing the style of works

2.3.1 Feature Extraction Methods for Film and Television Work Elements

(1) Algorithmic framework

The feature extraction method combining spatial pyramid matching and context-dependent histogram is divided into three steps: firstly, SIFT (Scale-invariant feature transform) features are extracted from each image and quantized into visual words; secondly, spatial partitioning of the image is carried out to construct spatial pyramids with different scale sizes; and lastly, the context-dependent histogram features are extracted and composed into feature vectors for each sub-block of the image. Finally, for each sub-block in the image, context-sensitive histogram features are extracted and composed into feature vectors.

(2) SIFT feature extraction and visual word quantization

Given a paper-cut image x, the original feature extraction algorithm combining pyramid modeling with context-dependent histograms consists of the following three steps.

First, SIFT features are extracted for each image and quantized into visual words. First, the local feature points of image x are captured by Dfference of gaussians detector and the 128-dimensional SIFT features of each feature point are extracted, and then the SIFT features in image x are quantized to get the visual words based on the visual lexicon consisting of k visual words obtained by pre-training through the K-means clustering algorithm.

(3) Constructing feature histograms

Spatial division of images constructs spatial pyramids with different scale sizes. The traditional image representation based on visual words ignores the spatial distribution relationship between visual words. And in the pyramid model, in order to get the spatial structural relationship that exists between these visual words, we sample the image to get l-layer pyramid.

Finally, the context-dependent histogram features are extracted separately for each sub-block obtained from the division in the image, and the original features are composed. Assume that the visual word spatial cooccurrence matrix within the subblock of the division unit is $C=\left[c_{ij}\right]\in R^{K\times K}$, where c_{ij} denotes the number of times that the visual word i appears adjacent to the visual word j with a distance less than a certain distance d within the subblock. In this way, the context-dependent histogram of the binary structure as in Eq. (15) can be defined based on the Markov steady state features:

$$S=\left[\pi^{a},\pi\right] \tag{15}$$

The K-dimensional vector in Eq. (15), which is the diagonal element of the spatial covariance matrix C of visual words, expresses the existence of spatial information between the same visual words; π is also a K-dimensional vector, whose value is computed according to Eq. (16):

$$\pi_i=\frac{\sum_{j\neq i}c_{ij}}{\sum_i\sum_{j\neq i}c_{ij}} \tag{16}$$

where π_i reflects the normalized value of the spatial distance between word i and other words in the image that is less than d. In this way, for the image X, it is divided into $\left(4^L-1\right)/3$ sub-blocks according to the pyramid model, and each sub-block is extracted from the $2K$-dimensional vectors S according to the spatial context-dependent histogram algorithm, and the vectors of the sub-blocks can be finally merged into the spatially-constrained R^{raw}-dimensional primitive features $V^{raw}=\left\{v_{mr}^{raw}\,|\,1\le m\le M,1\le r\le R^{raw}\right\}$, where $R^{raw}=2\times K\times\left(4^L-1\right)/3$.

2.3.2 Color feature extraction methods for film and television works

In the grid points distributed in the area of film and television work scenes, the characteristic spatial structure of auxiliary visual elements of film and television work scene color matching graphics is extracted, and combined with the three-dimensional distribution of characteristic quantities of film and television work scene color matching graphics, the characteristic equation is obtained as:

$$W^{ij}(x,y)=\frac{G^{ij}(x,y)}{\varepsilon}+\frac{n}{x^{ij}y^{ij}} \tag{17}$$

where W^{ij} is the amount of visual auxiliary features, $G^{ij}(x,y)$ is the central moment with $\left(x^{ij},y^{ij}\right)$ as the distribution of film and television work scene area.

Through the pixel marking, the edge contour marking in the process of color matching of film and television work scenes is carried out, and the contour line of the color matching of film and television work scenes is obtained as:

$$G^{ij}(x,y)=\frac{\sigma^2}{2^x-2x^{ij}+2y^{ij}} \tag{18}$$

Through the neighborhood feature matching method, the pixel point distribution matrix of the reconstructed movie and television work scene is obtained as:

$$\mathbf{Q}(x,y)=\begin{bmatrix} 0 & -\cos a & 1 \\ \cos(\theta) & -\sin(\theta) & 1 \\ \sin a & 1 & 0 \end{bmatrix} \tag{19}$$

where x-axis, y-axis and film and television work scene color distribution coordinate axis x',y' parallel, a for the animation scene pixels of the parameter of the horizontal coordinates of the value of, in each sub-region to calculate the film and television work scene color characteristics of the distribution of the wavelet parameter, to get the wavelet characteristics of the decomposition model for:

$$Q_w(x,y)=\lambda(w)Q_0(x,a)+\frac{1}{\lambda}(w)Q_0(y,b) \tag{20}$$

Among them:

$$Q_0(x,y)=\sigma_{xy}\frac{\sigma\overline{x}}{2x^2y^2}+\frac{\sigma_y}{\sigma_x} \tag{21}$$

$$\lambda(w)=\left(\frac{a}{w}+\frac{w}{b}\right)Q(ab) \tag{22}$$

where b is the value of the vertical coordinate of the parameter of the pixel of the animation scene, Q_0 is the value of the eigenvalue function, and $\lambda(w)$ is the value of the state parameter function.

Based on the feature decomposition model of each region, the visual color elements of the film and television works are extracted to provide data basis and model support for the image analysis of color matching vision.

2.3.3 Algorithm for extracting rhythmic features of dubbing for film and television works

(1) Note onset detection

The combination of energy and phase can effectively detect the onset of notes. The specific method is to use the complex domain information obtained by the fast Fourier transform as a combination of energy and phase can be a good combination of energy and phase changes in the relevant information. The calculation is shown in equation (23):

$$\tilde{S}_k(m) = \tilde{R}_k(m)e^{j\tilde{\phi}_k(m)} \tag{23}$$

where m is the number of the frame, $\tilde{R}_k(m)$ is the amplitude of the previous frame, and $\tilde{\Phi}_k(m)$ is computed from the phase difference between the previous frame and the one before it:

$$\tilde{\Phi}_k(m) = \text{princ}\ \ \arg\left[2\tilde{\varphi}_k(m-1) - \tilde{\varphi}_k(m-2)\right] \tag{24}$$

where princarg is the mapping of the obtained phase values to the $[-\pi, \pi]$ interval. The actual value of the k th frequency band is calculated as:

$$(m) = R_k(m)e^{j\varphi_k(m)} \tag{25}$$

where $R_k(m)$ and $\varphi_k(m)$ are the amplitude and phase of the current frame obtained after a short-time Fourier transform, respectively. The features of each frame are computed as:

$$\Gamma(m) = \sum\nolimits_{k=1}^{k=K} \left| S_k(m) - \tilde{S}_k(m) \right|^2 \tag{26}$$

By calculating the features of all frames of the audio and normalizing them, the note onset detection signal is obtained, which is a temporally continuous signal.

(2) Beat Period Estimation

Music rhythm has continuity and periodicity, so the beat period can be estimated from the start point detection signal. A better beat cycle estimation is the context-based beat cycle estimation algorithm. The first step of beat period estimation is to compute the autocorrelation function for the start detection function of each frame. In order to make the autocorrelation function more explicit, it is necessary to preprocess the data of each frame by first setting an adaptive moving average threshold, which is calculated as:

$$\bar{\Gamma}(m) = \sum\nolimits_{q=m-\theta}^{q=m+\theta} \Gamma_i(q) \tag{27}$$

The sliding window size is set to 16 points. Each point of the probe function is then allowed to subtract the corresponding threshold and half-wave rectify the output, calculated as in equation (28):

$$\tilde{\Gamma}_i(m) = \frac{\left(\Gamma_i(m) - \bar{\Gamma}_i(m)\right) + \left|\Gamma_i(m) - \bar{\Gamma}_i(m)\right|}{2} \tag{28}$$

Finally the autocorrelation function is calculated for the preprocessed signal as in equation (29):

$$A(i) = \frac{\sum_{m=1}^{N} \tilde{\Gamma}_i(m)\tilde{\Gamma}_i(m-i)}{|i-N|} \tag{29}$$

where $i = 1, 2, \cdots, N$ denotes the number of points on the frame and N denotes the frame length. The points τ_i in the autocorrelation domain can be mapped to the beat velocity with the mapping relation as in equation (30):

$$\text{BPM} = \frac{60}{\tau_i * 0.0116} \tag{30}$$

Since beats have continuity, the beat speeds of two neighboring frame arrays are intrinsically related, which can be taken into account when estimating the beat period of each frame of data. Based on the above considerations, the calculation of the beat rate t_b for the current frame can depend on the estimated beat rate t_{b-1} for the previous frame. A Hidden Markov Model can be constructed for this purpose. Each column of the state transfer matrix of this model consists of a Gaussian distribution with standard deviation 8. The state transfer matrix is calculated as in equation (31):

$$A\left(t_i, t_j\right) = P\left(t_b = t_j \mid t_{b-1} = t_i\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-\left(t_i - t_j\right)^2}{2\sigma^2}\right) \tag{31}$$

where σ is fixed to $8, t_i, t_j$ ranging from 0 to 127. The initial probability distribution is the Rayleigh distribution. The observation sequence is the autocorrelation function of each frame so that it can be decoded using the Viterbi algorithm. Firstly, the state probability vector of the current frame needs to be calculated, which is calculated by multiplying the state probability vector of the previous frame by the corresponding vector of the corresponding state transfer matrix and taking the maximum value. The calculation method is:

$$\Delta_b\left(t_j\right) = \max\left(\sum_{tj=0}^{tj=127} A\left(t_i, t_j\right)\Delta_{b-1}\left(t_j\right)\right) \tag{32}$$

The state probability for that frame is then multiplied by the autocorrelation function of the corresponding point in that frame, i.e:

$$\Delta_b\left(t_j\right) = \Delta_b\left(t_j\right) * A_b\left(t_j\right) \tag{33}$$

The velocity of the current frame is indexed by the maximum value of the state probability vector:

$$t_b = \arg\max\left(\Delta_b\left(t_j\right)\right) \tag{34}$$

The mapping of points to beat periods can be determined by equation (30).

(3) Beat tracking

Beat tracking can use a Hidden Markov Model to predict the exact point in time at which a beat occurs, based on the start point detection function. The beat tracking algorithm first assigns a state Φ to each point of the start point detection function, representing the distance of the point at that location from the previous beat point. This distance is expressed as a number of points. For example, if the tth point is a beat point, its state Φ_t is 0, and its next point's state Φ_{t-1} has a value of 1.

3 Case studies of film and television adaptations of classic Chinese and foreign literature

3.1 Structural Displacement in Film and Television Adaptations of Chinese and Foreign Classic Literature

3.1.1 Keyword Analysis of Chinese and Foreign Classic Literature Film and Television Works

Based on TF-IDF to extract the keywords of the movie and television works of Pride and Prejudice and Dream of the Red Chamber, and the first 15 keywords are shown in a chart, the results are shown in Table 1. In extracting the 15 keywords of the film and television work Pride and Prejudice, the overall expression of the plot to be embodied by the content of the film and television, in the literary novel Pride and Prejudice is mainly about the emotional entanglements between the heroine Elizabeth and the aristocrat Darcy, and the class error that exists between the two people, which are also shown to the audience through the movie; in the TV series Dream of Red Mansions, Baoyu and Daiyu, as the development of the story of the In the TV series “Dream of Red Mansions”, Baoyu and Daiyu as the main characters of the story promote the development of the plot. In the literary work “Dream of Red Mansions”, the author embodies the social status quo at that time according to the other supporting roles as well as reflects the tragedy of the women at the bottom of the society, which are also reflected in the TV series “Dream of Red Mansions”. It can be seen that the movie and television works adapted from Chinese and foreign literature can restore the characters' plots and artistic emotions in the literary content, and the overall integration effect is better.

Table 1: TF-IDF keywords

"Pride and Prejudice"		"A Dream in Red Mansions"	
1	Elizabeth	1	precious jade
2	Darcy	2	Daiyu
3	Pride	3	Grandma Jia
4	Prejudice	4	Baochai
5	Marriage	5	Wang Xifeng
6	Love	6	Showplace
7	Society	7	Jia Family
8	Fortune	8	Daiyu Buries Flowers
9	Money	9	Xiren
10	Estate	10	Empress's Visit
11	Entail	11	Qingwen
12	Ball	12	Poetry
13	Proposal	13	Search sb.'s house and confiscate his property
14	Elopement	14	Power
15	Reputation	15	Xiangling

3.1.2 Theme Analysis of Chinese and Foreign Classic Literary Works

LDA topic perplexity is a common metric for evaluating LDA topic models. It is a metric used in statistical modeling to evaluate the predictive effectiveness of a model, and is used to measure the predictive ability of an LDA model for unseen textual data.The smaller the LDA topic perplexity, the better the model performs in predicting unknown text. The perplexity will decrease with the increase of the number of topics, but too many topics will cause overfitting, so the number of topics at the local minimum of the perplexity curve should be taken as the optimal number of topics. The formula for calculating the perplexity is as follows:

$$Perplexity(D) = \exp\left(-\frac{\sum_{m=1}^{M} \log P\left(W_m\right)}{\sum_{m=1}^{M} N_m} \right) \tag{35}$$

The denominator in Eq. (35) is the sum of all words in the test set, i.e., the total length of the test set. Where P(W) refers to the probability of each word appearing in the test set. The topic perplexity curve is shown in Fig. 4, according to the principle that the smaller the perplexity indicator is locally the better, it can be seen by observation that the topic perplexity is locally the lowest when the number of topics is 7, so the optimal number of topics is 7. The resulting 7 topics and their keywords are visualized and represented.

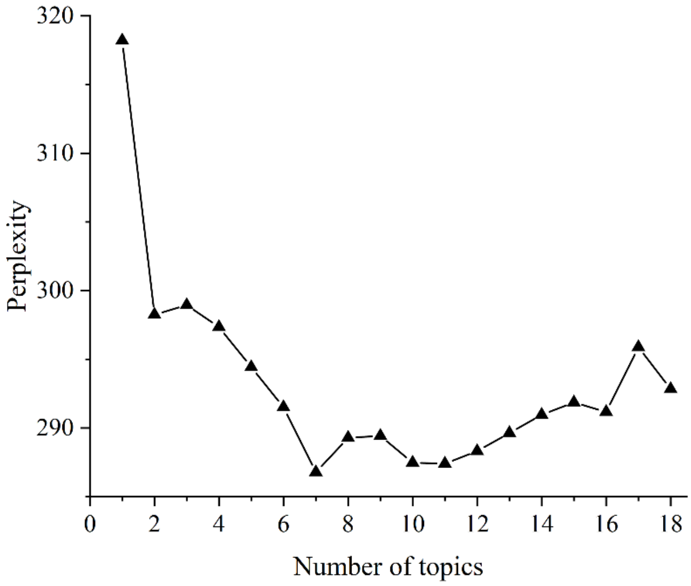
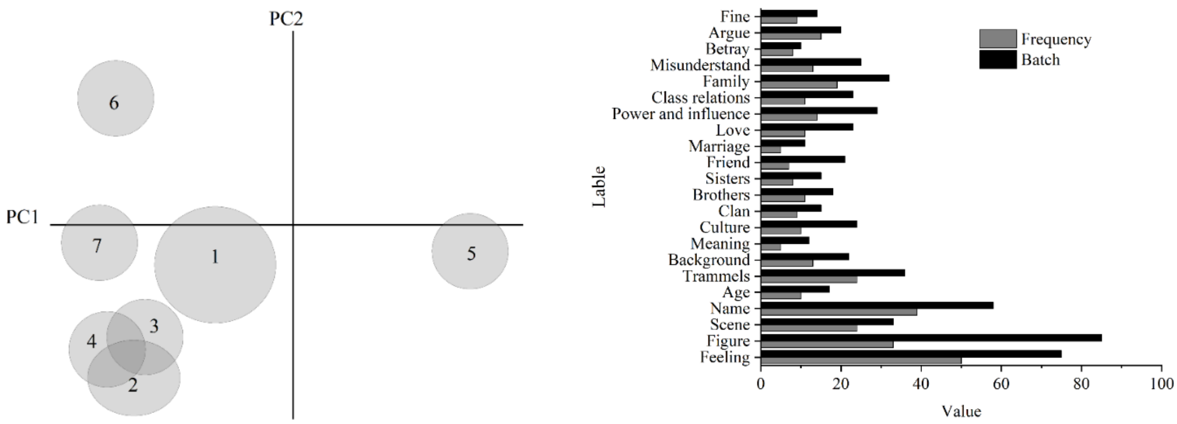


Figure 4: Topic confusion curve

Subsequently, theme visualization is carried out, and the results are shown in Fig. 5, (a) and (b) are the number of theme line-of-sight plots and theme vocabulary, respectively. The black bar in the figure represents the frequency of the word in the whole text, and the gray bar represents the batch of the word in the theme, and the number of batches of the keyword “character” in the figure is 75, which accounts for a relatively large proportion of the total number of batches. In literature, “character” is the main part of the whole text, which is the core element to establish the story, relationship, emotion and the direction of the whole text, and it is the trigger point to let readers have emotional resonance.



(a) Number of subjects at visual range (b) Thesaurus

Figure 5: Theme Visualization

The keywords for the specific seven themes are shown in Table 2.

Table 2: Number of optimal topics and their keywords

Theme	Keywords
1	Person, Era, Name, Background, Culture
2	Emotion, Meaning, Connection, Family
3	Brothers, Sisters, Friends, Parents, Lovers
4	Marriage, Love, Friendship, Family
5	Danger, Enemy, Trap, Conspiracy, Entrapment
6	Misunderstanding, Betrayal, Threats, Selling out, Arguing, Coercion
7	Beauty, Harmony, Wishes, Reunion, Harmony

3.2 Stylistic Integration in Film and Television Adaptations of Chinese and Foreign Classical Literature

3.2.1 Film color measurement

(1) Hue Measurement

Color is the most intuitive ideographic symbol of film, and the natural function and expressive value of color are integrated in film. Hue is an important color element. Through hue, movies can express different emotions, atmospheres and psychological activities. In this paper, Pride and Prejudice and A Dream of Red Mansions are calculated to find out the proportion of each color in the two films, and the specific results are shown in Figure 6. From a general point of view, the proportion of warm colors (red, yellow) shows a relatively obvious advantage, accounting for more than 20% more, but there are also some cold colors that account for more than warm colors, indicating that the film screen is dominated by warm colors.

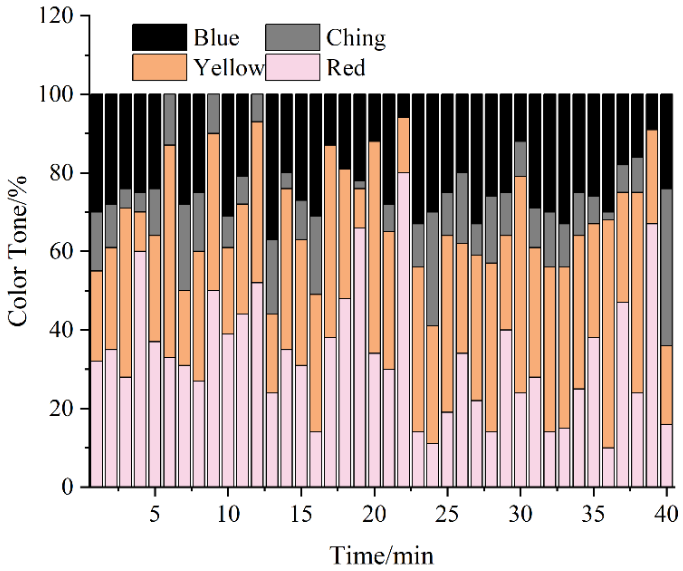


Figure 6: Distribution chart of the proportion of film and television colors

(2) Color saturation analysis

Color saturation is the degree of vividness of a color, also called purity, and is one of the three elements of the HSV color space model. Its value is a percentage, between 0-100%. The higher the saturation value, the more vivid the color, and vice versa. As the high saturation of the color is easy to make the human eye produce visual fatigue, in general, the color saturation value of about 50%, the human eye feels more comfortable. Figure 7 shows the average value of color saturation of two film and television works, and it can be found that the average saturation of colors in film and television works is generally not high, and its value is basically below 30%.

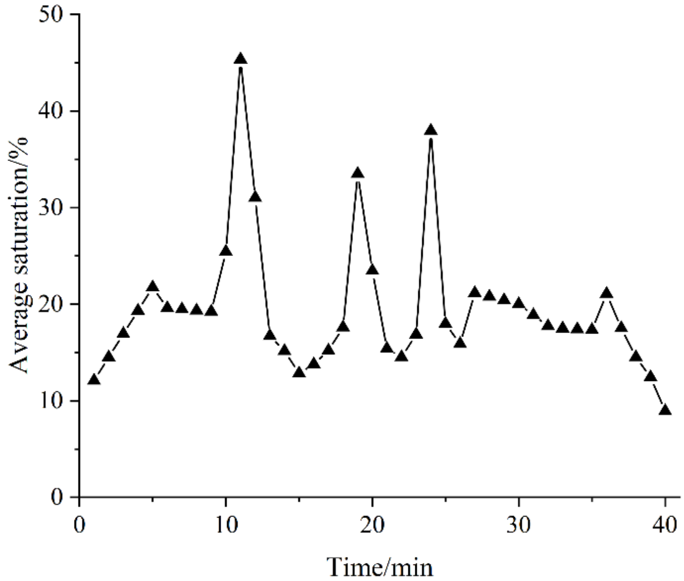


Figure 7: Average Saturation of Film and Television

3.2.2 Analysis of soundtracks for film and television productions

In order to verify the practicability of the automatic movie and TV extraction system designed in this paper, the movie and TV dubbing clips of four emotions, namely, panic, excitement, relaxation and sadness, were selected from two movies and TVs, namely, Pride and Prejudice and Dream of the Red Chamber, with 100 clips for each emotion. The system of this paper is used to extract the rhythmic characteristics of 400 movie and TV dubbing segments, obtain the BPM distribution graphs of different emotional dubbing, and describe the differences of different emotional dubbing. The histogram of the distribution of the rhythmic characteristics of dubbing with different emotions is shown in Figure 8. From the figure, the BPM peaks of panic and excitement are higher in the dubbing clips with different emotions, which indicates that when there are panic and excitement clips in the movie and television, the rhythmic characteristics of the dubbing are more significant and easy to be extracted. Percussion instruments are mostly used in the sound effects and background music production in the emotion clips of frightening and exciting. In contrast, the BPM peaks of the relaxed and sad mood clips are lower, indicating that the dubbed rhythmic features are not obvious and not easy to be extracted in such clips. For such clips, orchestral instruments are used more often.

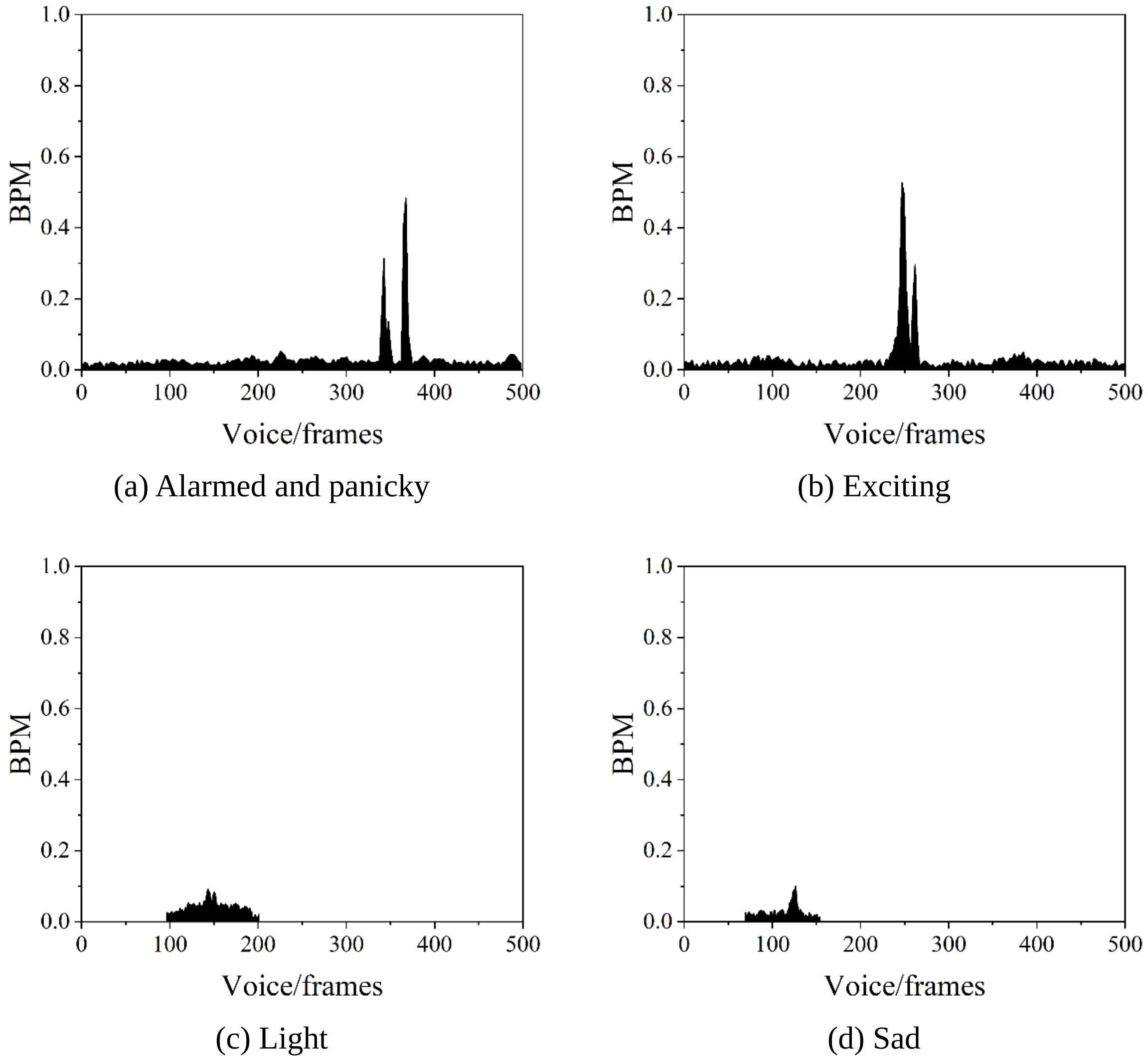


Figure 8: Rhythm distribution of different emotions

The system in this paper, the dubbing feature extraction system based on spectral energy distribution, and the feature extraction system based on tone-related fundamental frequency were used to classify the emotions of the selected 400 dubbing segments, respectively, and the results obtained are shown in Tables 3-Table 5. According to the statistical results of emotion classification in Table 3-Table 5, the accuracy, recall, and F1 value of the emotion classification results of the three different systems are determined, and the results are shown in Figure 9.

According to the emotion classification results in Table 3-Table 5 and Figure 9, the system in this paper extracts the rhythmic features of film and television dubbing for emotion classification, and the accuracy, recall, and F1 value are higher than those of the dubbing feature extraction system based on spectral energy distribution and the feature extraction system based on intonation-related fundamental frequency. The system based on tone-related fundamental frequency is better than the system based on spectral energy distribution for the classification of panic and excitement, indicating that the system is more effective for the classification of dubbing with higher BPM peaks. The differences in the classification detection indexes of this system for the four emotions are relatively smooth, and the classification effect is also better for the dubbing of relaxation and sadness with lower BPM peaks, which indicates that the system in this paper is able to accurately extract film and television rhythmic features, which is conducive to the classification of the emotions of film and television dubbing.

Table 3: Emotion classification results based on spectral energy distribution system

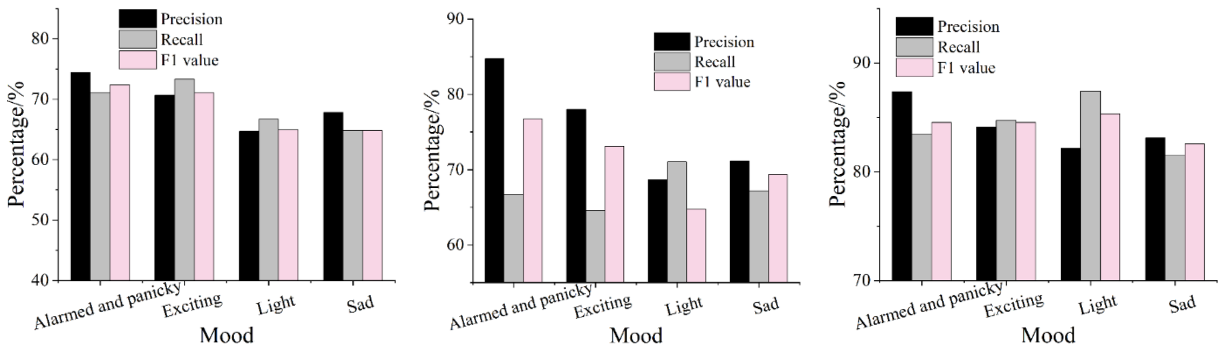
Mood	Alarmed and panicky	Exciting	Light	Sad
Alarmed and panicky	77	22	6	6
Exciting	17	54	10	8
Light	7	6	68	28
Sad	8	6	25	71

Table 4: Systematic emotion classification results based on pitch-related fundamental frequency

Mood	Alarmed and panicky	Exciting	Light	Sad
Alarmed and panicky	79	21	4	6
Exciting	16	76	11	9
Light	7	10	72	20
Sad	6	9	22	73

Table 5: The results of the systematic classification of emotions in this paper

Mood	Alarmed and panicky	Exciting	Light	Sad
Alarmed and panicky	89	10	5	4
Exciting	12	86	9	3
Light	5	7	85	12
Sad	4	5	14	85



(a) Frequency-dependent energy distribution (b) Tone-related fundamental frequency (c) System of this article

Figure 9: Comparison of the results of emotion classification in different systems

Comparing the resource occupancy of the three systems in the process of extracting movie and TV dubbing features and testing the energy consumption of this paper's system, the results are shown in Table 6. The CPU occupancy and memory occupancy of this paper's system are the lowest among the 3 systems, which are 3.46% and 1.42% respectively, indicating that this paper's system consumes less energy when extracting the rhythmic features of movie and TV dubbing.

Table 6: Resource usage of different systems

Feature extraction system	CPU occupancy rate/%	Memory usage/%
System of the text	3.46	1.42
Tone-related fundamental frequency system	8.06	3.51
A system based on spectral energy distribution	9.15	3.14

3.2.3 Visualization of the structure of adapted film and television

Based on the statistics obtained from the lens length, the number of shots and time data, produced a film and television lens length distribution and editing rate comparison as shown in Figure 10, film and television lens length of the various numerical values of the comparison is shown in Table 7. Comparing the waveform pattern of the movie in the figure, it can be seen at a glance that the distribution of movie and television shot length and editing rate have more obvious differences in different periods. In contrast, "Dream of the Red Chamber" has the largest numerical difference, which indicates that the difference in the length of individual shots in the movie is very large, combined with Figure 10, it is also not difficult to find out that the number of long shots in "Dream of the Red Chamber" is very large, which affects the normal distribution of the length of the whole movie shots.

Table 7: Compare the values of film and lens lengths

Title	Pride and Prejudice.1	Pride and Prejudice.2	A Dream in Red Mansions.1	A Dream in Red Mansions.2
Duration (minutes: seconds)	0:45:22	1:09:47	0:10:08	1:37:12
Number of lenses (pieces)	371	543	55	493
Average shot length (seconds)	8.9	8.3	12.6	12.7
Mid shot length(seconds)	5.9	5.3	11.5	9.2
Minimum shot length (seconds)	0.6	0.5	1.5	0.8
Maximum shot length(seconds)	104.4	45.2	33.8	96.7

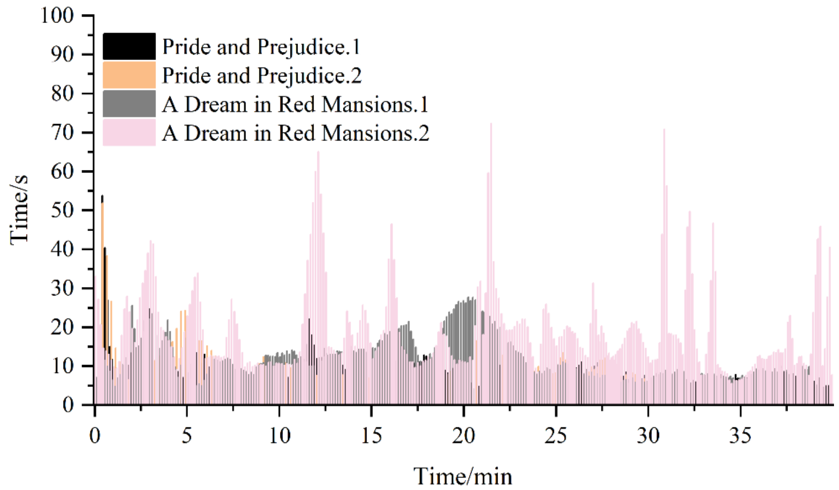
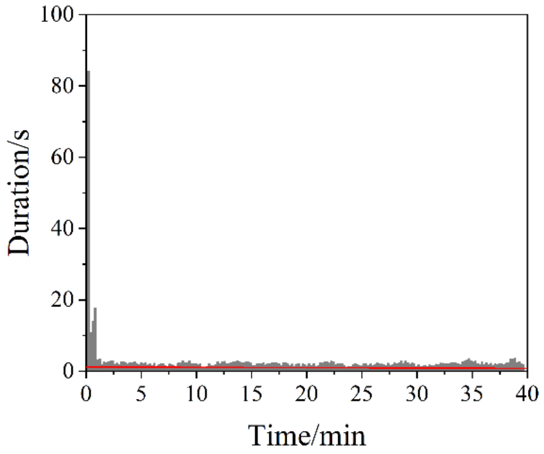
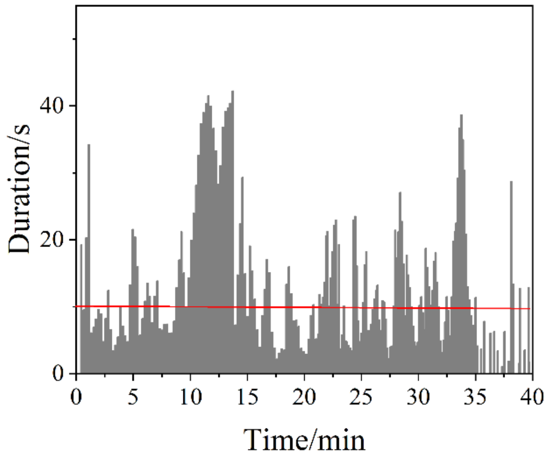


Figure 10: Film and television lens length distribution and editing rate comparison

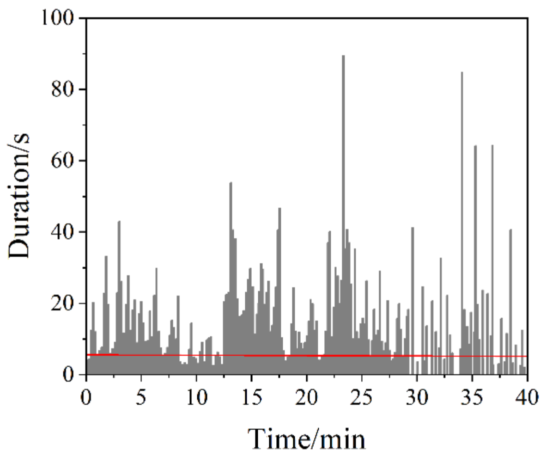
Rhythm is usually composed of editing intensity and movement within the frame, Pride and Prejudice and The Dream of the Red Chamber part of the clip shot length distribution dynamic curve shown in Figure 11, the red dynamic curve of the film's editing intensity of the ups and downs are not too much, the rhythm of the camera transitions are relatively smooth.



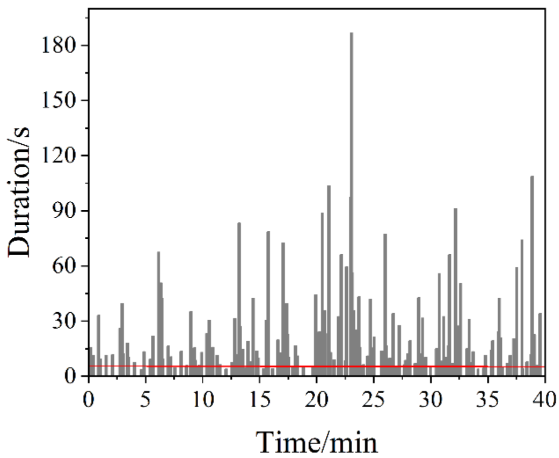
(a) Pride and Prejudice.1



(b) Pride and Prejudice.2



(c) A Dream in Red Mansions.1



(d) A Dream in Red Mansions.2

Figure 11: Dynamic curve of lens length distribution

3.3 Narrative Strategies and Story Content of Visualized Adaptations

3.3.1 Thematic changes and sublimation in visualized adaptations

Compared with literary works, movie and television works have a wider range of "potential readers". Therefore, the themes presented in film and television adaptations need to be in line with audience expectations and mainstream values. Popular themes that can evoke a sense of resonance and empathy are the first choice for most film and television adaptations. A safe popularized theme is safe, but it is difficult to further explore more "potential readers", so innovation becomes an effective way to open up the market. According to Yao Si, "If the meaning of a work is unexpected and beyond the horizon of expectation, it is exciting, and this new experience enriches and expands the new horizon of expectation."

3.3.2 Plot additions and deletions in visualized adaptations

Movie and TV adaptation works in the process of adaptation will often be part of the original works of the plot deletion, and according to the need to add appropriate content to assist the transformation of the movie and TV. This step is the key to the success or failure of the adaptation of a film or televisionized work. Plot patterns are divided into linear and non-linear in the context of narratology, with linear also known as storytelling and non-linear downplaying the story and characters. Most of the works of network novels focus on storytelling and take the linear plot mode, which is easy to be transformed into camera language, which is one of the reasons why network novels are favored by TV drama scriptwriters. However, the film and television adaptation cannot completely rely on the original text, a good screenwriter will add or delete the necessary plot by adjusting the original work, enriching the story, sublimating the main idea, and meeting the audience's expectations.

4 Conclusion

This paper takes the Chinese and foreign literature Pride and Prejudice and Dream of the Red Chamber as the research object, and uses TF-DIF, LDA theme model and deep learning to conduct metrological analysis, and on the basis of the work style feature extraction method based on spatial information, the combination of image features combined with spatial pyramid model and contextual correlation histogram is used to construct the work style model. The results of the study show that 15 high-frequency words are extracted based on TF-DIF, and 7 themes of Chinese and foreign literary works are shown through visualization. The color styles of the two films and TV shows are mainly in warm tones, orchestral instruments are used more, and the rhythms of camera transitions are relatively smooth in both cases.

Acknowledgements

Li Fan, research on Narrative Transformation in the Adaptation of Online Novels into Film and Television Dramas, Hebei Province Cultural and Artistic Science Planning and Tourism Research Project, 2024 (Project No. HB24-YB137).

About the Author

Li Fan (1980.3-), female, the Han nationality, from Zhangjiakou, Hebei Province, Master's degree, Associate Professor, Research Direction: Theory of Film and Television Literature.

Yongfei Ye (1978.6-),female, the Han nationality, from Zhangjiakou, Hebei Province, Master's degree, Professor, Research Direction: Smart Education.

References

[1] Mehmood, I., Azeem, M., Javaid, A., & Khuram, M. (2023). Globalization and English Literature: Explaining how globalization has affected the themes, style, and influences of modern English literature. Jahan-e-Tahqeeq, 6(4), 25-33.

[2] He, D., Li, W., & Li, Y. (2021, May). Analysis of the Path of Spreading Red Culture by Film and TV Works in the New Era. In 7th International Conference on Humanities and Social Science Research (ICHSSR 2021) (pp. 789-792). Atlantis Press.

[3] Xiang, L. (2022). Analysis on the media presentation and dissemination path of cultural variety shows. Frontiers in Art Research, 4(11).

[4] Chen, Y., & Zhuang, G. (2022, March). Problems and Countermeasures of Film and Television Adaptation of Literary Works in the New Era. In 8th International Conference on Education, Language, Art and Inter-cultural Communication (ICELAIC 2021) (pp. 348-352). Atlantis Press.

[5] Diko, M., & Saule, N. (2024). Adapting a Novel for Television: A Systemic Critique. International Journal of African Renaissance Studies-Multi-, Inter-and Transdisciplinarity, 19(1), 27-43.

[6] Pasopati, R. U., Suyaji, F. I. M., Kirana, K. S., Ramadhani, R. D., & Wijaya, K. (2024). The Intricateness of Adaptation of Literature to Film in Today's Crisscrossed World. Journal Corner of Education, Linguistics, and Literature, 3(4), 390-400.

[7] Gillespie, D. (2014). The Art of Literary Adaptation and English-Language Film Interpretations of Russian Literature ('Anna Karenina'). Procedia-Social and Behavioral Sciences, 154, 30-35.

[8] Arkan, G. N. (2023). Cinema Literary Adaptations as a Narrative Form. Contemporary Issues of Communication, 2(2), 50-58.

[9] Primorac, A. (2017). Victorian literature and film adaptation. Victorian Literature and Culture, 45(2), 451-459.

[10] Graffy, J. (2011). Literature and film. The Cambridge Companion to Twentieth-Century Russian Literature, 235-250

[11] Rahmoun, O. (2020). Film Adaptation between the Pride of Literature and the Prejudice of Inferiority. AWEJ for Translation & Literary Studies, 4.

[12] Lanqi, Z. (2023). Analysis of Film and Television Adaptation of Classical Literary Works. Frontiers in Art Research, 5(11).

[13] Chua, B., & Ho, E. (2023). Introduction: Global literary adaptation in the twenty-first century. In The Routledge Companion to global literary adaptation in the Twenty-First Century (pp. 1-15). Routledge.