Correlation Feature Selection Analysis for Fault Diagnosis of Induction Motors

Chapter · December 2016

DOI: 10.1007/978-3-319-24584-3_104

5 authors, including:

Thanaporn Likitjarernkul
Prince of Songkla University

Rakkrit Duangsoithong
Prince of Songkla University

Kusumal Chalermyanont
Prince of Songkla University

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Advanced Computer and Communication Engineering Technology	
Series Title		
Chapter Title	Correlation Feature Selection Analysis for Fault Diagnosis of Induction Motors	
Copyright Year	2016	
Copyright HolderName	Springer International Publishing Switzerland	
Corresponding Author	Family Name	**Likitjarernkul**
	Particle	
	Given Name	**Thanaporn**
	Prefix	
	Suffix	
	Division	Faculty of Engineering, Department of Electrical Engineering
	Organization	Prince of Songkla University
	Address	Hat Yai, Songkhla, 90112, Thailand
	Email	l.thanapond@gmail.com
Author	Family Name	**Sengchaui**
	Particle	
	Given Name	**Kiattisak**
	Prefix	
	Suffix	
	Division	Faculty of Engineering, Department of Electrical Engineering
	Organization	Prince of Songkla University
	Address	Hat Yai, Songkhla, 90112, Thailand
	Email	ak.kiattisak@hotmail.com
Author	Family Name	**Duangsoithong**
	Particle	
	Given Name	**Rakkrit**
	Prefix	
	Suffix	
	Division	Faculty of Engineering, Department of Electrical Engineering
	Organization	Prince of Songkla University
	Address	Hat Yai, Songkhla, 90112, Thailand
	Email	rakkrit.d@psu.ac.th
Author	Family Name	**Chalermyanont**
	Particle	
	Given Name	**Kusumal**
	Prefix	
	Suffix	
	Division	Faculty of Engineering, Department of Electrical Engineering
	Organization	Prince of Songkla University
	Address	Hat Yai, Songkhla, 90112, Thailand

	Email	kusumal.c@psu.ac.th
Author	Family Name	**Prasertsit**
	Particle	
	Given Name	**Anuwat**
	Prefix	
	Suffix	
	Division	Faculty of Engineering, Department of Electrical Engineering
	Organization	Prince of Songkla University
	Address	Hat Yai, Songkhla, 90112, Thailand
	Email	anuwat.p@psu.ac.th

Abstract	This paper presents a feature selection method for stator winding fault analysis of induction motors by using a Correlation-based Feature Selection (CFS) method. The 14 original motor parameters are selected from the feature selection method with various searching approaches. The classification efficiency of optimal features obtained from the feature selection method is compared with results from the feature extraction method and the original features. In our experiment, we employ a 2.2 kW delta-connected motor which drives a dc generator as a load. The experimental results demonstrate that 4 common selected features for stator winding fault analysis of induction motors are a percent of load (*%Load*), a power factor (*pf*), a negative sequence voltage (V_n), and a negative sequence impedance (Z_n). The accuracy of the classification using this feature subset is higher than using all original features for three classification methods.

Keywords (separated by '-')	Feature selection - Induction motor - Stator winding faults

Correlation Feature Selection Analysis for Fault Diagnosis of Induction Motors

Thanaporn Likitjarernkul, Kiattisak Sengchaui, Rakkrit Duangsoithong, Kusumal Chalermyanont and Anuwat Prasertsit

Abstract This paper presents a feature selection method for stator winding fault analysis of induction motors by using a Correlation-based Feature Selection (CFS) method. The 14 original motor parameters are selected from the feature selection method with various searching approaches. The classification efficiency of optimal features obtained from the feature selection method is compared with results from the feature extraction method and the original features. In our experiment, we employ a 2.2 kW delta-connected motor which drives a dc generator as a load. The experimental results demonstrate that 4 common selected features for stator winding fault analysis of induction motors are a percent of load (*%Load*), a power factor (*pf*), a negative sequence voltage (V_n), and a negative sequence impedance (Z_n). The accuracy of the classification using this feature subset is higher than using all original features for three classification methods.

Keywords Feature selection · Induction motor · Stator winding faults

T. Likitjarernkul (✉) · K. Sengchaui · R. Duangsoithong · K. Chalermyanont · A. Prasertsit
Faculty of Engineering, Department of Electrical Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112, Thailand
e-mail: l.thanapond@gmail.com

K. Sengchaui
e-mail: ak.kiattisak@hotmail.com

R. Duangsoithong
e-mail: rakkrit.d@psu.ac.th

K. Chalermyanont
e-mail: kusumal.c@psu.ac.th

A. Prasertsit
e-mail: anuwat.p@psu.ac.th

H.A. Sulaiman et al. (eds.), *Advanced Computer and Communication Engineering Technology*, Lecture Notes in Electrical Engineering 362,
DOI 10.1007/978-3-319-24584-3_104

1 Introduction

Electrical motors are critical equipments for any machine in process industries. Motor deteriorations or any fault occurred since rotors and stators have been forced by electromagnetic field all the time. Moreover, mechanical damages caused by thermal and electrical stresses have an effect to the performance and the lifetime of motors.

About 37 % [1] of induction motor faults are stator winding faults due to the deterioration of the winding insulation from contamination of oil, humidity, and sewage. They impact on opening or shorting one or more circuits of windings.

Current and voltage signals of the induction motor contain information of stator winding faults, and they are widely used to detect and locate stator winding faults in various methods. The Motor Current Signature Analysis (MCSA) method is one of the most frequently used methods to analyze the motor fault by identifying stator current spectrums in abnormal harmonics [2, 3]. In addition, the Extended Park's Vector Approach (EPVA) is applied to analyze EPVA signatures by identifying a spectral component at twice the fundamental supply frequency [4]. Motor sequence components (i.e. negative and zero sequence components of the current) are also used for stator fault diagnosis [5–8]. Moreover, modeling and simulation studies can provide useful information about the electric behavior of the motors, and they relate to the analysis of the presence of the internal fault in the stator windings [9, 10]. Other techniques, such as an instantaneous angular speed technique, temperature monitoring, air-gap torque monitoring, magnetic flux monitoring, noise/acoustic noise, induced voltage monitoring, surge testing, gas analysis, and partial discharge [11] are also used in order to diagnose stator winding faults.

Currently, some or all electrical features of induction motors mentioned previously are used for the motor fault analysis and detection. However, using all or inappropriate electrical features will increase the complexity of the system and the stodgy storages Moreover, they may not be able to classify the motor faults correctly. Consequently, the feature selection should be required in as preprocessing in order to reduce original features and extract the appropriate features. Basically, there are two methods to reduce the feature dimension. One is the feature selection and the other is feature extraction. The feature selection can be used to choose optimal features from original features to remove irrelevant and redundant of original features and also decreases the complexity of the system. The feature extraction is another method to reduce a number of features by transforming original features to lower dimensional spaces. The Principal Component Analysis (PCA) [12] is one of the example of the feature extraction. The PCA reduces the dimension of features without eliminating the signal information using the principal component. The proposed fault diagnosis system is shown in Fig. 1.

This paper presents the use of the correlation feature selection for selecting electric features. Optimal electric features are obtained by the feature selection for stator winding fault analysis of induction motors. The accuracy of motor fault classification obtained by the optimal features from the feature selection method is

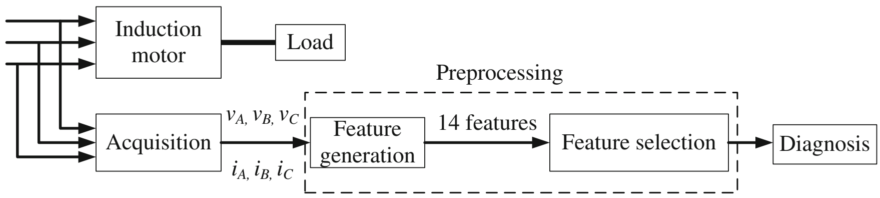

Fig. 1 The proposed fault diagnosis system

compared with one that obtained from the feature extraction method and original features. Various theories related to electrical features calculation are reviewed in Sect. 2. In Sect. 3, the feature selection methods are explained. Section 4, an experimental setup is described and the experimental results of the feature selection analysis are showed and discussed in Sect. 5 and conclude in Sect. 6.

2 Feature Generation

The electrical features were calculated by motor current and voltage signals. These features are popular features for the stator winding fault diagnosis. These signals were fed to the preprocessing block to reduce a number of features, and optimal features are searched and selected for such faults. The original features in this paper can be obtained from the following methods.

2.1 Symmetrical Components

The positive and the negative sequence components of the induction motor are normally used to indicate the stator faults. Figure 2 shows a workflow for extracting the positive (I_p) and the negative (I_n) sequence components from three-phase

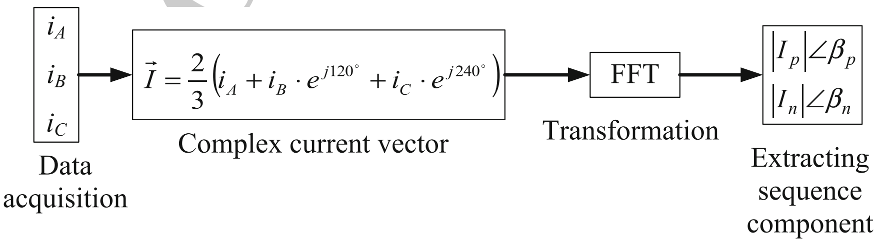

Fig. 2 Signal processing workflow to extract sequence components

Layout: T1 Standard Unicode | Book ID: 352772_1_En | Book ISBN: 978-3-319-24582-9
Chapter No.: 104 | Date: 16-10-2015 | Time: 8:15 pm | Page: 4/10

current signals. The three-phase currents are used to construct a complex current vector ($\overrightarrow{I}$), as expressed by

$$\overrightarrow{I} = \frac{2}{3}\left(i_A + i_B \cdot e^{j120^\circ} + i_C \cdot e^{j240^\circ}\right) \tag{1}$$

where i_A, i_B, and i_C are the currents in the phase A, B, and C, respectively.

Fast Fourier Transform (FFT) of the complex current vector can automatically separate the positive and the negative sequence currents for all frequency components. An example of an actual current spectrum is shown in Fig. 3. The three-phase voltages are processed in the similar approach.

2.2 Extended Park's Vector Approach (EPVA)

The EPVA method is the observation of the spectrum of the Park's vector module. The motor current Park's vector components (i_D, i_Q) are

$$i_D = \left(\sqrt{\frac{2}{3}}\right) i_A - \left(\sqrt{\frac{1}{6}}\right) i_B - \left(\sqrt{\frac{1}{6}}\right) i_C \tag{2}$$

$$i_Q = \left(\sqrt{\frac{1}{2}}\right) i_B - \left(\sqrt{\frac{1}{2}}\right) i_C \tag{3}$$

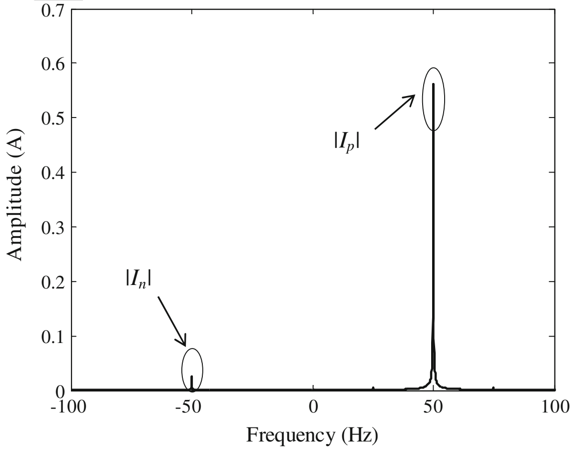

Fig. 3 A magnitude spectrum of the complex current vector

Table 1 The original features

	Features	Description
1	*%Load*	Percent of load
2	*pf*	Power factor
3	I_p	Positive sequence current
4	*Angle* (I_p)	Angle of positive sequence current
5	I_n	Negative sequence current
6	*Angle* (I_n)	Angle of negative sequence current
7	V_p	Positive sequence voltage
8	*Angle* (V_p)	Angle of positive sequence voltage
9	V_n	Negative sequence voltage
10	*Angle* (V_n)	Angle of negative sequence voltage
11	Z_p	Positive sequence impedance
12	Z_n	Negative sequence impedance
13	I_{dc}	Magnitude of square of Park's vector module at DC level
14	$I_{100\mathrm{Hz}}$	Magnitude of square of Park's vector module at twice the supply frequency

The square of the Park's vector module is given by

$$|i_D + ji_Q|^2 = \left(\frac{3}{2}\right)\left(i_d^2 + i_i^2\right) + 3i_d i_i \cos(2\omega t - \alpha_d - \beta_i) \tag{4}$$

where i_d is the maximum value of the direct sequence current, i_i is the maximum value of the reverse sequence current, ω is the angular frequency (rad/s), t is the time variable (s), α_d is the initial phase angle of the direct sequence current (rad), and β_i is the initial phase angle of the reverse sequence current (rad).

The square of the Park's vector module can be used to identify unbalanced three-phase currents. The spectrums of dc level and the component located at twice the supply frequency are obtained byapplying the FFT to the square of the Park's vector module [1].

Two feature generation methods contain 12 original features and two extra features including the percent of load (*%Load*) and the power factor (*pf*) are added to be the original features. The 14 original features as shown in Table 1 are used to be the inputs for the next feature selection process.

3 Feature Selection

The Feature selection is a process to select optimal features from original features. It can reduce a number of features by removing irrelevant and redundant features. Basically, it can be divided into four categories: Filter, Wrapper, Hybrid, and Embedded methods [13]. The Filter method is the feature selection that applied independent evaluation criteria without involving any classification algorithms with measurement technique such as CFS [14] and consistency based subset evaluation. The Wrapper method applies a classification algorithm for a feature subset evaluation. This method is better than the Filter method, but it takes longer time for a computation. The Hybrid method combines the advantage of above two approaches. Finally, the Embedded method has built-in the feature selection in classifier.

Generally, the feature selection has four steps [15]: subset generation, subset evaluation, stopping criteria, and result validation. First, it searches optimal features by using searching algorithms. Then, this subset is evaluated by subset evaluator, and it stop by stopping criteria. Finally, it validates selected features.

Searching algorithms for finding the feature set have several methods which are shown below

1. **Best first** is the searching method that selects the feature with the best heuristic value.
2. **Exhaustive search** searches all possible feature subset.
3. **Greedy stepwise** is the searching method. It starts with empty or full feature set. Then, it adds the suitable feature or removes the inappropriate feature.
4. **Linear forward selection** (**LFW selection**) is the searching method that begins with an empty set and successively adding features.
5. **Random search** randomly selects the feature subset from original features.
6. **Rank search** selects the feature subset from ranking of total features.

CFS [14] is a well-known feature selection method that considers the correlation between features and classes and between features and other features. Relevance of the feature subset can be defined by using Pearson's correlation equation [13] which is expressed by (5)

$$Merit_s = \frac{kr_{kc}}{\sqrt{k + (k-1)r_{kk}}} \quad (5)$$

where k is the number of features, c is the number of classes, $Merit_s$ is relevance of the feature subset, r_{kc} is the average linear correlation coefficient between these features and classes, and r_{kk} is the average linear correlation coefficient between different features.

The linear correlation coefficient is defined by

$$r = \frac{\sum_i (x_i - \overline{x_i})(y_i - \overline{y_i})}{\sqrt{\sum_i (x_i - \overline{x_i})^2}\sqrt{\sum_i (y_i - \overline{y_i})^2}} \quad (6)$$

where i is the number of values (x or y), x_i is the x value for observation i, $\overline{x_i}$ is the mean x value, y_i is the y value for observation i, and $\overline{y_i}$ is the mean y value.

Layout: T1 Standard Unicode | Book ID: 352772_1_En | Book ISBN: 978-3-319-24582-9
Chapter No.: 104 | Date: 16-10-2015 Time: 8:15 pm | Page: 7/10

4 Experimental Setup

The three-phase four-pole delta-connected induction motor is used in the experimental setup as shown in Fig. 4. The motor parameters and ratings are summarized in Table 2. The motor is modified for interturn stator winding faults in each phase. A shorting resistor is used to limit the short-circuit current in the winding not exceed to 5 A. The induction motor is monitored by three current sensors and three voltage sensors. The measured signals are sent to the computer through a National Instruments (NI) data acquisition device with 6000 Hz sampling rate and ten operate conditions. The operation conditions contain of four cases: healthy motor, and the short- turns motor of 7, 15, and 31 turns in each phase. Each condition is operated under 6, 30, 60, and 90 % rated load of the motor. Note that the 14 original features are obtained from the normalized signals with rated parameters of the testing motor. Classes for the fault classification consist of '0', 'A', 'B', and 'C' as the following meanings

'0' is Healthy motor
'A' is Interturn fault in phase A
'B' is Interturn fault in phase B
'C' is Interturn fault in phase C

Optimal features analyzed from the correlation feature selection algorithms are used in three well-known classifiers: k-Nearest Neighbor or kNN ($k = 3$), Naïve Bayes, and Decision Tree. The accuracy rated of classification will be compared with one obtained from the feature extraction using the PCA by using data set of 1320 samples. The result is validated by using 10 folds cross-validation technique.

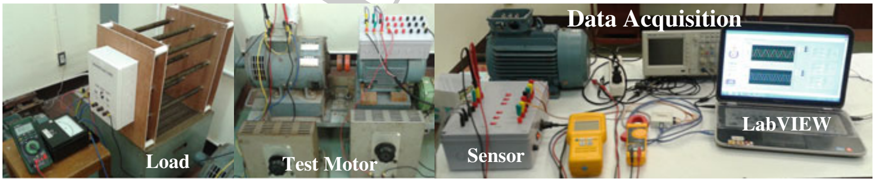

Fig. 4 Experimental setup

Table 2 Parameters and ratings of test machines

V	Hz	r/min	kW	cosØ	A
230Δ/400Y	50	1430	2.2	0.79	8.66/4.98
415Y	50	1435	2.2	0.765	4.94

5 Results and Discussions

The original features are analyzed by CFS for the feature selection method with aforementioned search algorithms. The results in Table 3 show that the correlation feature selection can effectively select 4 or 5 optimal features, while the PCA can reduce to 8 features. The 4 common features including *%Load, pf,* V_n, and Z_n are chosen from each search algorithms. These features are the part of optimal features for stator winding fault analysis. It is shown that these features have high correlation between features and classes, and they have low correlation between features and other features.

According to experimental results, the accuracy rates using original features are 58.5606, 43.0303, and 85.0758 % for kNN, Naïve Bayes, and Decision Tree, respectively. Based on feature selection, the accuracy rates using Decision Tree is higher than using kNN and Naïve Bayes. These classifiers provide the accuracy rates as ranged 85–87, 75–78, and approximately 43 %, respectively. Significantly, it can be explained that using the feature selection to select optimal features can reduce a number of features and increase the accuracy of the classification system. Fault classification provides better performance than using the feature extraction and original features, respectively.

Table 3 Classification correction of the proposed approach

Search algorithms	Subset evaluators	Number of selected features	Selected features	Accuracy rate (%)		
				kNN ($k = 3$)	Naïve Bayes	Decision tree
Original features	–	14	Original feature set	58.5606	43.0303	85.0758
Best first	CFS	5	*%Load, pf,* I_n, V_n, Z_n	75.1515	43.7121	86.0606
Exhaustive search	CFS	5	*%Load, pf,* I_n, V_n, Z_n	75.1515	43.7121	86.0606
Greedy stepwise	CFS	4	*%Load, pf,* V_n, Z_n	76.1364	43.0303	85.6818
LFW selection	CFS	5	*%Load, pf,* I_n, V_n, Z_n	75.1515	43.7121	86.0606
Random search	CFS	4	*%Load, pf,* V_n, Z_n	78.1818	42.803	87.1212
Rank search	CFS	4	*%Load, pf,* V_n, Z_n	77.0455	42.9545	85.6061
Ranker	PCA	8	–	60.5303	43.1061	59.8485

6 Conclusions

This paper presents the correlation feature selection for stator winding fault analysis of the induction motor. According to our experimental results, it can be found that the common selected features for stator winding fault analysis of the induction motor are *%Load, pf,* V_n, and Z_n. These features are good indicators to predict stator winding faults, and they can be applied for any size of motors. Furthermore, using the feature selection and the feature extraction can improve the accuracy of the classification system. For future works of our research, more features will be considered and compared with other feature selection methods. AQ1

Acknowledgments This work was fully supported by Prince of Songkla University Graduate Studies Grant Contract number 95000201 and partially supported by Center of Excellence in Wireless Sensor Networks (CoE-WSN), Faculty of Engineering, Prince of Songkla University, Hat Yai, Songkhla, Thailand.

References

1. Bonaldi, E.L., de Oliveira, L.E.L., Borges da Silva, J.G., Lambert-Torres, G., Borges da Silva, L.E.: Predictive maintenance by electrical signature analysis to induction motors. In: Araujo R.E (ed.) Induction Motors Modelling and Control, InTech (2012)
2. Jung, J.-H., Lee, J.-J., Kwon, B.-H.: Online diagnosis of induction motors using MCSA. IEEE Trans. Industr. Electron. **53**(6), 1842–1852 (2006)
3. Thomson, W.T.: On-line MCSA to diagnose shorted turns in low voltage stator windings of 3-phase induction motors prior to failure. In: IEEE International Conference on Electric Machines and Drives Conference, IEMDC 2001, pp. 891–898 (2001)
4. Cruz, S.M.A., Cardoso, A.J.M.: Stator winding fault diagnosis in three-phase synchronous and asynchronous motor, by the extended park's vector approach. IEEE Trans. Industr. Appl. **37**(5), 1227–1233 (2001)
5. Ben Khader Bouzid, M., Champenois, G.: New expressions of symmetrical components of the induction motor under stator faults. IEEE Trans. Industr. Electron. **60**(9), 4093–4102 (2013)
6. Arkan, M., Perovic, D.K., Unsworth, P.: Online stator fault diagnosis in induction motors. Proc. IEE Electr Power Appl. **148**(6), 537–547 (2001)
7. Cheng, S., Habetler, T.G.: A new method to detect stator turn to turn faults in a closed-loop multiple-motor drive system. In: IEEE international symposium on diagnostics for electric machines, power electronics and drives, SDEMPED 2009, pp. 1–6 (2009)
8. Cheng, S., Zhang, P., Habetler, T.G.: An impedance identification approach to sensitive detection and location of stator turn-to-turn Faults in a closed-loop multiple-motor drive. IEEE Trans. Industr. Electron. **58**(5), 1545–1554 (2011)
9. Chang, X., Cocquempot, V., Christophe, C.: A model of asynchronous machines for stator fault detection and isolation. IEEE Trans. Industr. Electron. **50**(3), 578–584 (2003)
10. Bachir, S., Tnani, S., Trigeassou, J.-C., Champenois, G.: Diagnosis by parameter estimation of stator and rotor faults occurring in induction machines. IEEE Trans. Industr. Electron. **53**(3), 963–973 (2006)
11. Siddique, A., Yadava, G.S., Singh, B.: A review of stator fault monitoring techniques of induction motor. IEEE Trans. Energy Convers. **20**(1), 106–114 (2005)
12. Jolliffe, I.T.: Principal Component Analysis. Springer, New York (1986)

13. Duangsoithong, R., Windeatt, T.: Relevance and redundancy analysis for ensemble classifiers. In: P. Perner (ed.) Machine Learning and Data Mining in Pattern Recognition, MLDM 2009, LNAI 5632, pp. 206–220. Springer, Heidelberg (2009)
14. Hall, M.A.: Correlation-based feature selection for discrete and numeric class machine learning. In: Proceeding of the 17th International Conference on Machine Learning, pp. 359–366. Morgan Kaufmann, San Francisco (2000)
15. Liu, H., Yu, L.: Toward integrating feature selection algorithms for classification and clustering. IEEE Trans. Knowl. Data Eng. **17**(4), 491–502 (2005)

Author Query Form

Book ID :	**352772_1_En**
Chapter No.:	**104**

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author
During the process of typesetting your chapter, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query Refs.	Details Required	Author's Response
AQ1	No queries.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	,
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	