Analytical Methods

Accepted Manuscript

Analytical Methods

www.rsc.org/methods

Volume 8 | Number 1 | 7 January 2016 | Pages 1–224

ISSN 1759-9660

ROYAL SOCIETY OF CHEMISTRY

PAPER
Tetsuo Okada et al
Chiral resolution with frozen aqueous amino acids

175 YEARS

ROYAL SOCIETY OF CHEMISTRY

rsc.li/methods

1 The Provenance of Nephrite in China Based on Multi-spectral Imaging

2 Technology and Gray-Level Co-occurrence Matrix

3 Dian Chen [a, b], Ming Pan[c], Wei Huang[d], Wugan Luo[a, b*], Changsui Wang[a, b]

4 a. Key Laboratory of Vertebrate Evolution and Human Origin, Chinese Academy of Sciences. Institute of
5 Vertebrate Paleontology and Paleoanthropology, Beijing 100049, China

6 b. Department of Archaeometry, University of Chinese Academy of Sciences, Beijing 100049, China

7 c. Graduate School, People's Public Security University of China, Beijing 100038, China

8 d. Institute of Forensic Science, Ministry of Public Security, National Engineering Laboratory for Forensic Science,
9 Beijing 100038, China

10

11 **Abstract:** The provenance of nephrite is the basis for the research, restoration and collection of
12 nephrite artifacts. However, it is not easy to determine provenance, because the major, minor and
13 trace elements in nephrite from different sources usually overlap one another. In recent years,
14 submicro-structural methods were applied to probe into the provenance of nephrite. However, the
15 submicro-structural imaging of nephrite could not be well preserved and the image characteristics
16 could not easily be described quantitatively. On the basis of such situation, this paper first
17 introduces non-destructive multi-spectral imaging technology, which can be used to directly obtain
18 and record the structural images. And then the gray-level co-occurrence matrix is adopted to
19 establish a set of parameters to describe the characteristics of nephrite. Finally, multivariate
20 statistics methods including principal component analysis and hierarchical clustering are used to
21 achieve a visual testing result. This study shows that the method is a simple and effective approach
22 to classify a limited number of samples and suggests that the method could be applied to the
23 research on the provenance of nephrite artifacts after further improvement.

24 **Keywords:** Nephrite; Provenance; Multi-spectral imaging; Gray-level co-occurrence matrix;
25 Submicro-structure

26

27 **1. Introduction**

28 　　Nephrite is a monomineralic rock in general, mainly composed of fine grained tremolite
29 $[Ca_2Mg_5Si_8O_{22}(OH)_2]$ or actinolite $[Ca_2(Mg, Fe^{2+})_5Si_8O_{22}(OH)_2]$. Jade culture, particularly
30 nephrite culture, occupied a significant role in Chinese civilization. The use of nephrite as ancient
31 tools, ornaments and ritual objects in China can be traced back to 8000 years ago [1]. Since the
32 late Neolithic period, greater numbers of ancient nephrite artifacts were excavated from tombs and

*Corresponding author：Wugan Luo, Department of Archaeometry, University of Chinese Academy of Sciences,
Beijing 100049, China. Tel. :+86 10 15001161179; Fax: +86 10 88256501.E-mail addresses: xiahua@ucas.edu.cn

33   ruins, such as Hongshan culture, Liangzhu culture and Yin Ruins [2]. However, determining the

34   provenance of nephrite artifacts has remained one of the unsolved problems.

35   Previously, scholars tried to answer this question mainly from four aspects. First, morphology

36   analysis, which generally uses petrographic and scanning electron microscopes [3,4]. Second,

37   structural analysis by use of X-ray diffractometer, transmission electron microscope and Raman

38   spectroscope, etc. [5,6,7]. Third, elemental analysis including major, minor and trace elements

39   [8,9]. Fourth, isotopes ratio, such as lead, strontium and neodymium isotopes etc. [10]. However,

40   the mineralogical and petrological characteristics of nephrite from different provenances are

41   basically the same. There are nuances in structure caused by the crystallinity of the tremolite, such

42   as the crystal size, shape, and combination of the mineral particles. But that are not related to the

43   specific provenance. Methods such as isotopic analysis could reveal the age of the nephrite

44   mineral deposit formation, but it will cause damage to the nephrite artifacts. Finally, due to the

45   inhomogeneity distribution of impurities in nephrite, it is impossible to get convincing results by

46   using non-destructive examination of elements.

47   As is generally believed among many nephrite lovers, the analysis of nephrite provenance

48   using instruments is not as accurate as the connoisseurs' experience. The identification of nephrite

49   provenance in China, especially in jade trade, mainly relies on visual experience. Some

50   connoisseurs and collectors with rich experience affirm that the naked eyes can distinguish

51   different nephrite provenance based on the 'floccules structure' (more than 100μm) which is

52   caused by coarse particles and aggregates of tremolite. However, the correct rate is about 70% [11].

53   The 'floccules structure', named as 'submicro-structure' in this paper, is clearer than microscale,

54   but it is difficult to be preserved.

55   In this paper, we will discuss that the visual characteristics of submicro-structure can be

56   properly recorded by multi-spectral imaging technology. Furthermore, gray-level co-occurrence

57   matrix can give an quantitative description. At last, the results can be obtained by using

58   multivariate statistical analysis.

59

2. Materials and methods

61   *2.1. Nephrite samples*

62   There are about a dozen nephrite sources in China. In this experiment, 24 nephrite samples

63   were collected from Chinese deposits including Yecheng, Hetian, Qiemo, Ruoqiang in Xinjiang

64   province, Geermu in Qinghai province. Four other nephrite samples were from Chuncheon in

65   Korea and Baikal in Russia. The localities are shown in Fig. 1.

66

67 **Fig. 1** The geographic positions of main nephrite deposits in China, South Korea and Russia (solid circles refer to
68 the samples of this article).

69

70 These present nephrite samples from mountain areas have the same occurrence, which are
71 sequentially numbered 1 to 28. Among them, 1~6 is for Ruoqiang, 7~9 for Qiemo, 10~11 for
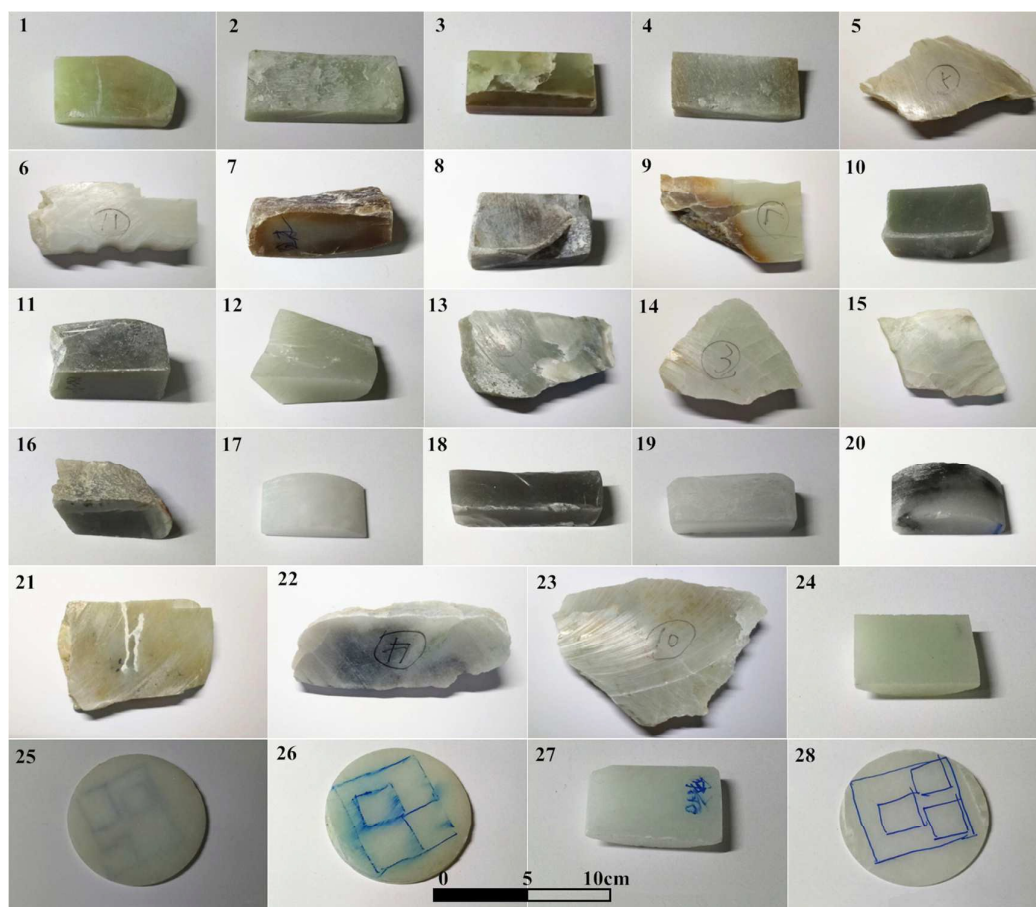72 Yecheng, 12~15 for Hetian, 16~23 for Geermu, 24~25 for Chuncheon, and 26~28 for Baikal (Fig.
73 2).

Fig. 2 Nephrite sample pictures analyzed in this article.

2.2. multi-spectral imaging

As mentioned above, the 'floccules structure' can be observed inside the nephrite under strong light. However, the submicro-structure is usually randomly distributed in the interior of the nephrite and it is a cumulative and compositive visual effect. Therefore, this situation seriously disturbs the observation of the naked eyes and the acquisition of image. But the multi-spectral imaging can directly show the reflection spectra of the submicro-structure as grayscale image. It can not only contain more information than chromaticity space but also effectively avoid the phenomenon of metamerism [12].

The multi-spectral imaging technology helps us to acquire multi-channel images of target samples by multiple-bandpass filters [13,14]. In particular, the incident full band or wide optical signal is divided into several narrow band beams, by which the target objects are irradiated and then images of different spectral bands will be generated. Similarly, the absorption of light in different wavelengths is not the same. So each bunch of light with a specific wavelength is reflected by the object, the light intensity will produce different changes, hence each channel will form one image with high spectral resolution. Furthermore, the discrete images obtained by

92　individual light wavelengths can synthetize a composite image by various of preprocessings such

93　as calibration, noise reduction, time registration, spatial registration and resampling. However, the

94　synthesized image is not gray but chromatic. It can overcome the limitations and differences of

95　single image in spectrum and spatial resolution as well as strengthen the performance of detail,

96　thus more comprehensive and reliable descriptions for the same target can be achieved [15].

97　Multi-spectral imaging has been applied to the field of art conservation and art history since

98　the early 1990s [16]. Convincing results have been achieved in material analysis and identification,

99　preservation status assessment, digital image archiving, and so on [17,18]. However, this

100　technology has not been used to explore the source of nephrite. Not only can multi-spectral

101　imaging reveal more information than the naked eyes and ordinary camera, but also the structure

102　of nephrite by use of a suitable light source, on account of the advantages of working mechanism

103　and its successful application in other fields..

104　*2.3. Background on Gray-Level Co-occurrence Matrix (GLCM)*

105　The submicro-structure in the nephrite can be viewed as a special texture. In figure analysis,

106　texture is a common concept of describing images. An image texture is a set of metrics calculated

107　from image processing. Image texture tells us the information about the spatial arrangement of

108　color or intensities in an image or selected region of an image [18]. Therefore, texture features

109　have been widely used for classification and recognition of images.

110　Since the GLCM were proposed by Haralick et al. in 1973, it has been utilized as the main

111　tool in image texture analysis. Haralick suggested that equations based on the co-occurrence

112　matrix could describe the image texture. It is a statistical method to characterize image texture

113　structure by comparing the gray level of the sample pattern to its surroundings. GLCM is

114　represented by the function:

115
$$P(i, j, d, \theta) = \{[(x, y), (x + Dx, y + Dy) \mid f(x, y) = i; f(x + Dx, y + Dy) = j]\}$$

116　Where x, y=1, 2,···, N-1, N are pixel coordinates of images and Dx, Dy are position offsets;

117　while i, j=1, 2, ···L-1, L are gray levels, in which i represents the gray level at location of

118　coordinate (x, y), and j represents the gray level of its neighboring pixel at a distance d and a

119　direction θ from a location (x, y).

120　Haralick [20] defined some characteristic parameters of GLCM for texture analysis. We used

121　the following nine textural features in this study after some experimenting. Let P (i, j, d, θ) is the

122　(i, j)th entry in a normalized GLCM.

123　1). Angular Second Moment:

124
$$A = \sum_{i=1}^{L} \sum_{j=1}^{L} p^2(i, j, d, \theta)$$

125　Sum of squares of every element in GLCM, also called Energy. It reflects the degree of
126　uniformity of grayscale distribution and roughness of images.

127   2). Entropy:

128     $$E = \sum_{i=1}^{L} \sum_{j=1}^{L} - p(i,\ j, d, \theta) \times \log p(i,\ j, d, \theta)$$

129   It is the measurement of the randomness of the image content. Representing the amount of
130   information of the image and indicating the complexity of the texture.

131   3). Moment of Inertia:

132     $$I = \sum_{i=1}^{L} \sum_{j=1}^{L} (i - j)^2 \times p(i,\ j, d, \theta)$$

133   It can largen the differences in spatial distribution of image grayscale and pick out the
134   complexity of the distribution.

135   4). Correlation:

136     $$C = \sum_{i=1}^{L} \sum_{j=1}^{L} (i \times j \times p(i, j, d, \theta) - u_1 \times u_2) / d_1^{\,2} \times d_2^{\,2}$$

137   Among which

138     $$u_1 = \sum_{i=1}^{L} i \sum_{j=1}^{L} p(i,\ j, d, \theta)$$

139     $$u_2 = \sum_{j=1}^{L} j \sum_{i=1}^{L} p(i,\ j, d, \theta)$$

140     $$d_1 = \sum_{i=1}^{L} (i - u_1)^2 \sum_{j=1}^{L} p(i,\ j, d, \theta)$$

141     $$d_2 = \sum_{j=1}^{L} (j - u_2)^2 \sum_{i=1}^{L} p(i,\ j, d, \theta).$$

142   Correlation shows the similarity in the direction of row or column of the elements in GLCM.
143   When the value of the matrix element is equal, the correlation will be large; on the contrary, while
144   the value of the matrix element is very different, the correlation should be small. If there is a
145   horizontal texture in the image, the correlation value of this direction matrix is greater than the
146   correlation values of the rest of the matrix. Therefore, correlation can be used to determine the
147   direction of the texture.

148   5). Inverse Difference Moment:

149     $$IDM = \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{p(i,\ j, d, \theta)}{1 + (i - j)^2}$$

150   IDM can reveal the homogeneity of image texture and measure the extent of the local change
151   in the texture of the image. If the diagonal elements are larger values, IDM will take the   larger
152   value. It shows there is a lack of difference between different regions of the image texture and the
153   local part is very homogeneous. Accordingly, the continuous grayscale image will have a larger
154   IDM value.

155   6). Contrast:

156     $$CON = \sum_{n=0}^{L-1} n^2 \{ \sum_{|i-j|=1} p(i,\ j, d, \theta) \}$$

157   It indicates the contrast of the brightness of a pixel and its neighborhood pixel in an image. If
158 elements deviated from the diagonal are larger values, i.e. image brightness values vary quickly,
159 CON will take the larger value. The greater the contrast is, the clearer the visual effect is.

160   7). Variance:

161
$$V = \sum_{i=1}^{L} \sum_{j=1}^{L} (i - u)^2 \times p(i,\ j, d, \theta)$$

162   Among which $u = \sum_{i=1}^{L} \sum_{j=1}^{L} \dfrac{p(i,\ j, d, \theta)}{L^2}$ is the mean.

163   It reflects the overall uniformity of the image.

164   8). Sum of Average:

165
$$SOA = \sum_{k=2}^{2L} k \times G(k)$$

166   Among which

167
$$G(k) = \sum_{i+j=k}^{L} \sum^{L} p(i,\ j, d, \theta)\ ,\quad k=2, \cdots, 2L.$$

168   Sum of average can show the light and shade on the image.

169   9). Sum of Variance:

170
$$VOA = \sum_{k=2}^{2L} (k - AOV)^2 \times G(k)$$

171   Among which

172
$$G(k) = \sum_{i+j=k}^{L} \sum^{L} p(i,\ j, d, \theta)\ ,\quad k=2, \cdots, 2L.$$

173   Sum of Variance can reflect the period of the texture.

174

175   **3. Measurement**

176   *3.1. Instrument for measurement*

177   The instrument used in this measurement is *CRi Nuance* multi-spectral imaging system (Fig.
178 3). The *Nuance Imaging Module* contains the principal imaging components in a single compact
179 enclosure: high-resolution scientific-grade CCD imaging sensor, solid-state LCTF with a polarizer,
180 wavelength tuning element, spectrally optimized lens and internal optics. The camera lens (Nikon
181 AF-S VR 105mm f/2.8G camera IF-ED, AF lens type S) is apochromatic in the range of
182 315nm~1100nm. A 16-band light source emitter (Australian Polilight PL 500) with high output
183 stability and power adjustable is used in darkroom condition to avoid interference of
184 miscellaneous light.

185   The images were acquired in the wavelength range 450-950nm with the scanning step ~10nm,
186 which cover the visible (VIS) and near infrared (NIR). They were stored in a laptop that was
187 cable-connected to the camera. Subsequent processing of the images was performed by means of

188   the *CRi Nuance* and *Matlab* programs.
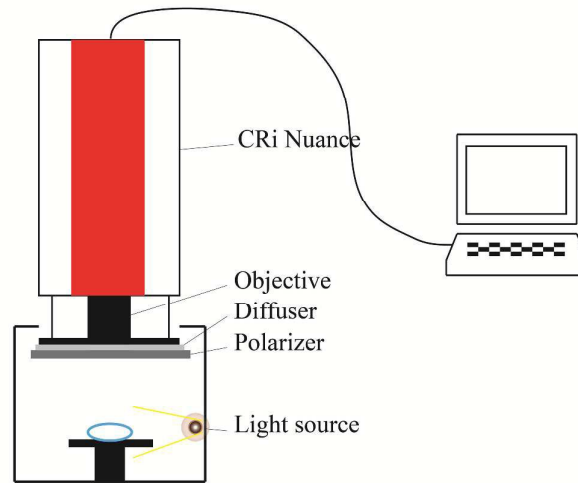
189

190   **Fig. 3** The experimental system.

191

192   *3.2. Measuring process*

193   In order to avoid the surrounding rock and defective parts, a marker was made on the
194   corresponding surface of the sample. Two areas without defects of each sample were selected with
195   the aim of verifying the uniformity and improving the data size at the same time. Then the sample
196   was placed on the carrier platform at the proper height and angle for the sake of the marked area in
197   the field of lens view. In the dark room, a beam of light at the angle of 45 degrees illuminated the
198   sample. Adjusting the focal length of the lens until the submicro-structure image of the sample
199   was clearly displayed. The exposure parameters were already set by the software to record and
200   save with the image.

201   Software *CRi Nuance 2.4* can synthesize an integrated image fused by the multi-spectral
202   images of each sample in different wavelengths. We used the synthesized image to represent the
203   sample on account of the reasons stated earlier (see *2.2.*). In this way, each sample had two
204   respective synthetic images based on different selected areas. To unify the standard, the region of
205    $1 \times 1 cm^2$ in focused position of every synthesized image would be further analysized and
206   processed.

207   .The gray level is very sensitive in the GLCM algorithm. Compressing the gray level
208   properly will greatly reduce the amount of computation, save the storage space and also depress
209   the noise effect in the image. But if the gray level is reduced too much, it may destroy the
210   distribution characteristics of the texture and remove the valuable features in the image [21].
211   Through test and comparison, we decided to compress the 256 gray levels in the original image to
212   8 gray levels. The software *Matlab 2015B* was used to process data. In particular, the step size was
213   set up as 1, the 9 characteristic parameters of each sample's GLCM in four directions of 0, 45, 90

214 and 135 degrees were obtained. For statistics, the characteristic parameters of four directions were

215 averaged (see Table 1). Then the statistical software *SPSS 20.0* was used to cluster analysis of the

216 data.

217

218 **Table 1**

219 Provenance and GLCM'S characteristic parameters of nephrite samples.

Number	Provenance	A	E	I	C	IDM	CON	V	SOA	VOA
1a	Ruoqiang	0.1390	2.1405	0.0359	0.2920	0.9820	0.0228	4.1725	8.3667	1.9629
1b	Ruoqiang	0.1373	2.1238	0.0324	0.2830	0.9829	0.0242	4.1682	8.8347	1.8932
2a	Ruoqiang	0.1416	2.1004	0.0256	0.2839	0.9872	0.0168	4.7737	9.5705	2.6914
2b	Ruoqiang	0.1453	2.1143	0.0214	0.2798	0.9912	0.0184	4.7824	9.2748	2.8084
3a	Ruoqiang	0.1287	2.2647	0.0783	0.2991	0.9608	0.0435	5.2036	10.4298	3.1178
3b	Ruoqiang	0.1302	2.2492	0.0821	0.2992	0.9724	0.0491	5.2020	10.2834	3.1284
4a	Ruoqiang	0.1298	2.1867	0.0396	0.2518	0.9802	0.0242	3.8185	7.6594	1.9325
4b	Ruoqiang	0.1278	2.1890	0.0421	0.2641	0.9730	0.0265	3.9021	7.8472	1.9748
5a	Ruoqiang	0.1361	2.1759	0.0461	0.2987	0.9770	0.0291	4.2325	8.4843	1.9941
5b	Ruoqiang	0.1403	2.1493	0.0458	0.2894	0.9698	0.0294	4.2814	8.3470	1.9988
6a	Ruoqiang	0.1530	2.0395	0.0293	0.3569	0.9854	0.0062	4.6705	9.3892	1.9031
6b	Ruoqiang	0.1483	2.0532	0.0384	0.3483	0.9873	0.0067	4.7201	9.2743	1.9247
7a	Qiemo	0.1408	2.1532	0.0429	0.3146	0.9785	0.0261	4.6513	9.3245	2.0602
7b	Qiemo	0.1385	2.1684	0.0238	0.3297	0.9804	0.0284	4.2983	9.4031	2.1084
8a	Qiemo	0.1394	2.1010	0.0209	0.2739	0.9895	0.0143	3.9804	7.9843	2.5796
8b	Qiemo	0.1425	2.1328	0.0238	0.2694	0.9834	0.0133	4.0821	8.0348	2.6843
9a	Qiemo	0.1272	2.2797	0.0683	0.2708	0.9658	0.0391	4.3700	8.7616	3.4283
9b	Qiemo	0.1299	2.2843	0.0792	0.2801	0.9673	0.0402	4.4972	8.7434	3.3583
10a	Yecheng	0.1520	2.0731	0.0317	0.3485	0.9841	0.0202	4.7907	9.6038	2.1004
10b	Yecheng	0.1593	2.0749	0.0289	0.3281	0.9833	0.0230	4.8210	9.7391	2.1948
11a	Yecheng	0.1257	2.2975	0.0750	0.2490	0.9625	0.0469	4.1900	8.3926	2.2839
11b	Yecheng	0.1243	2.2840	0.0824	0.2384	0.9683	0.0453	4.2821	8.3480	2.4834

12a	Hetian	0.1669	1.9729	0.0277	0.3312	0.9862	0.0197	4.0659	8.1514	2.6011
12b	Hetian	0.1634	2.0314	0.0293	0.3382	0.9871	0.0187	4.0729	8.3419	2.6484
13a	Hetian	0.1357	2.1374	0.0296	0.2603	0.9852	0.0178	4.6754	9.3759	3.9489
13b	Hetian	0.1298	2.1293	0.0320	0.2831	0.9857	0.0202	4.7921	9.4973	3.8947
14a	Hetian	0.1486	2.2204	0.0785	0.3058	0.9607	0.0428	3.7117	7.4474	2.1962
14b	Hetian	0.1502	2.2483	0.0802	0.3182	0.9589	0.0482	3.8514	7.3871	2.2283
15a	Hetian	0.1690	2.0706	0.0868	0.4933	0.9566	0.0449	3.3545	6.7374	2.0892
15b	Hetian	0.1602	2.1083	0.0875	0.4682	0.9489	0.0465	3.6502	6.6871	2.1273
16a	Geermu	0.1403	2.1158	0.0248	0.2460	0.9876	0.0180	3.7045	7.4290	3.1733
16b	Geermu	0.1468	2.1308	0.0259	0.2378	0.9805	0.0201	3.8014	7.5981	3.4752
17a	Geermu	0.1969	1.9120	0.0730	0.6351	0.9635	0.0383	3.8794	7.7866	0.3196
17b	Geermu	0.1834	2.0814	0.0684	0.6294	0.9670	0.0342	3.8915	7.8343	0.4573
18a	Geermu	0.1602	2.0327	0.0336	0.3760	0.9832	0.0157	3.9369	7.9073	0.8950
18b	Geermu	0.1724	2.0248	0.0402	0.3728	0.9820	0.0153	4.0124	8.0384	0.8745
19a	Geermu	0.1522	2.0610	0.0325	0.3582	0.9837	0.0219	4.4862	8.9924	1.3044
19b	Geermu	0.1545	2.0834	0.0349	0.3720	0.9741	0.0233	4.3932	9.0834	1.4073
20a	Geermu	0.1551	2.0924	0.0502	0.3906	0.9749	0.0301	4.3689	8.7593	1.0198
20b	Geermu	0.1593	2.1081	0.0533	0.3911	0.9780	0.0329	4.5492	8.8347	1.1284
21a	Geermu	0.1565	2.0717	0.0434	0.3819	0.9783	0.0262	4.4385	8.8992	1.1377
21b	Geermu	0.1548	2.0848	0.0483	0.3762	0.9739	0.0258	4.4037	8.9834	1.2849
22a	Geermu	0.1867	1.9383	0.0355	0.4815	0.9822	0.0200	4.1906	8.4081	0.5041
22b	Geermu	0.1934	1.9822	0.0382	0.4671	0.9731	0.0123	4.2047	8.2474	0.5733
23a	Geermu	0.1524	2.0970	0.0340	0.2698	0.9830	0.0240	3.5291	7.0754	2.5566
23b	Geermu	0.1611	2.1037	0.0375	0.2872	0.9803	0.0243	3.4749	7.0873	2.3845
24a	Chuncheon	0.1913	1.9745	0.0932	0.6301	0.9534	0.0488	3.3402	6.7073	0.0825
24b	Chuncheon	0.1839	1.9824	0.0992	0.6391	0.9604	0.0503	3.3821	6.9732	0.0889
25a	Chuncheon	0.1690	2.0706	0.0868	0.4933	0.9566	0.0449	3.3545	6.7374	0.0892
25b	Chuncheon	0.1794	2.1249	0.0892	0.4728	0.9693	0.0453	3.3756	6.8357	0.0843

26a	Baikal	0.1881	2.0159	0.1009	0.5874	0.9496	0.0511	5.2846	10.5991	0.4016
26b	Baikal	0.1832	2.0483	0.1323	0.5282	0.9573	0.0531	5.3813	10.3474	0.4024
27a	Baikal	0.1848	1.9516	0.1044	0.5672	0.9728	0.0543	5.1397	9.2966	0.3367
27b	Baikal	0.1924	2.0246	0.1123	0.5382	0.9689	0.0538	5.2490	9.7480	0.3723
28a	Baikal	0.1685	2.1606	0.1403	0.5221	0.9301	0.0671	5.1679	10.3730	0.4009
28b	Baikal	0.1714	2.1840	0.1370	0.5512	0.9542	0.0647	5.1734	10.2475	0.4038

220

4. Results and discussion

4.1. Grayscale images

In general, there are several kinds of visual images. Among them, the samples from Chuncheon show the most obvious characteristics (Fig. 4). The deep black parts are impurities which can cause some interference to the observation of submicro-structure which is composed of some bright spots and light gray patches. Compared with other samples, its submicro-structure is very loose. On the whole, it looks unclean and in disorder.
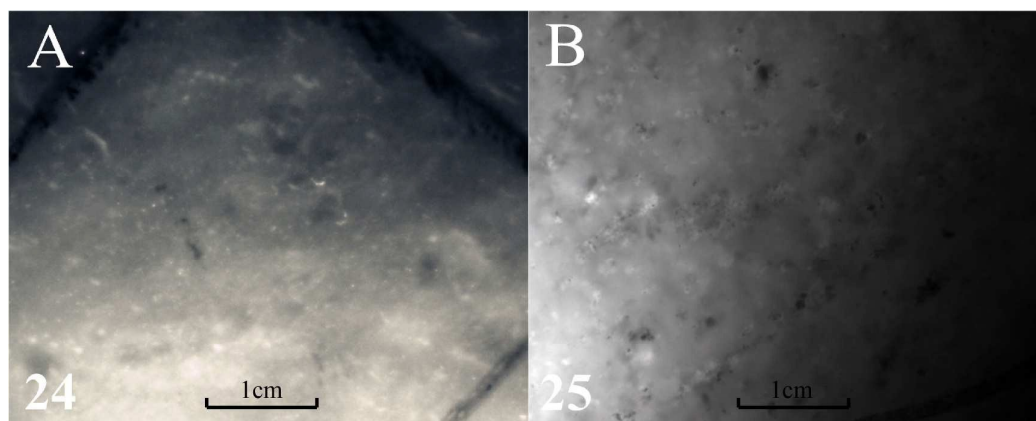
Fig. 4 The original gray pictures of nephrite samples from Chuncheon. The number of the lower left corner is the sample number and the rough black lines in the left-hand image is made with a marker pen.

Another distinctive group is the samples from Baikal (Fig. 5). There are many local areas that exceed the scale of $1cm^2$ on the photo of its submicro-structure. The whole is whiter and brighter, which is just like boiled rice porridge.
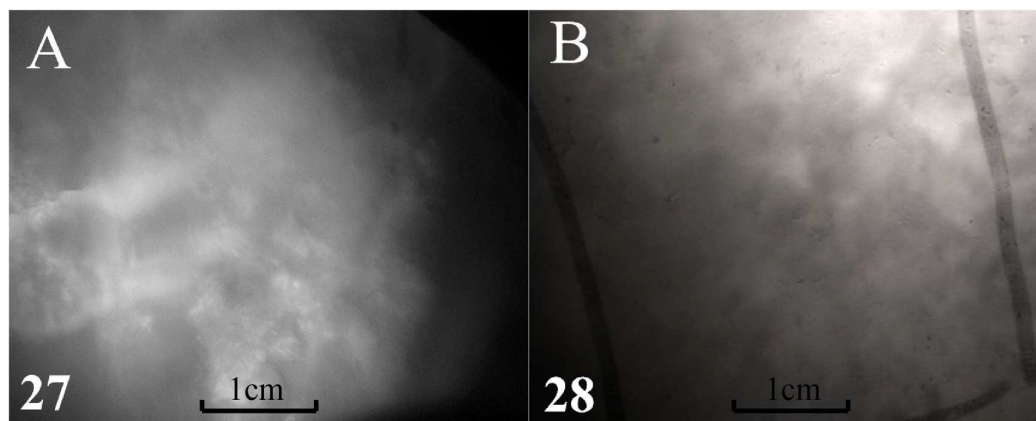
235

Fig. 5 The selected original gray pictures of nephrite samples from Baikal. The black lines in the right-hand image
is made with a marker pen.

236
237

238

239   The submicro-structures of typical Qinghai nephrite will appear particle-like texture (Fig. 6).
240 It is finer and smoother than nephrite from Chuncheon. Those white and bright spots are
241 especially like granulated sugar and have a regular distribution. Sometimes this part of the
242 structure can be observed directly with the naked eye under the conditions of natural light. So for
243 connoisseurs, it becomes a distinct feature from the nephrite of other sources [11].

244
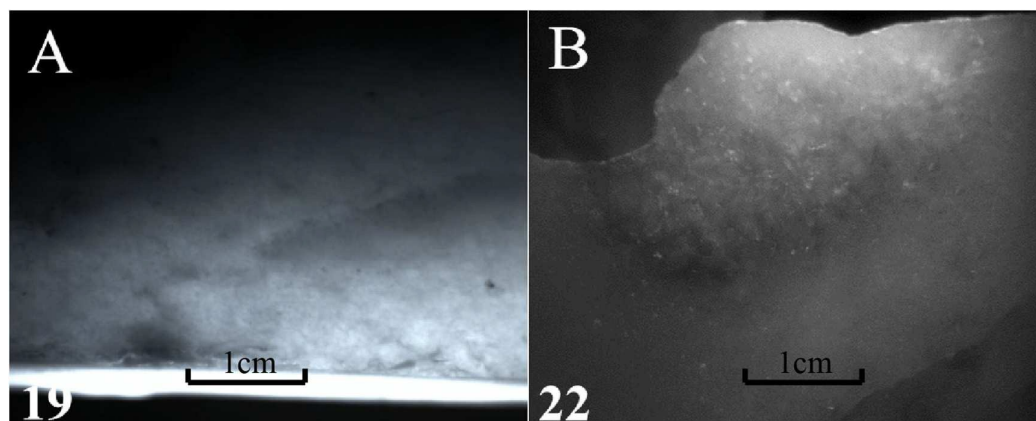
245

246   **Fig. 6** The selected original gray pictures of nephrite samples from Geermu, Qinghai.

247

248   There is no uniform submicro-structure in Xinjiang nephrite. Some even have no obvious
249 structure, such as Fig. 7A. Moreover, some have very delicate submicro-structure where black and
250 white spots interlace each other densely (Fig. 7B, D). But there are some special cases, such as Fig.
251 7C. A series of parallel thin white lines appear in the interior of the nephrite and they are not flaws
252 or fractures. This is indeed a kind of submicro-structure.

Analytical Methods
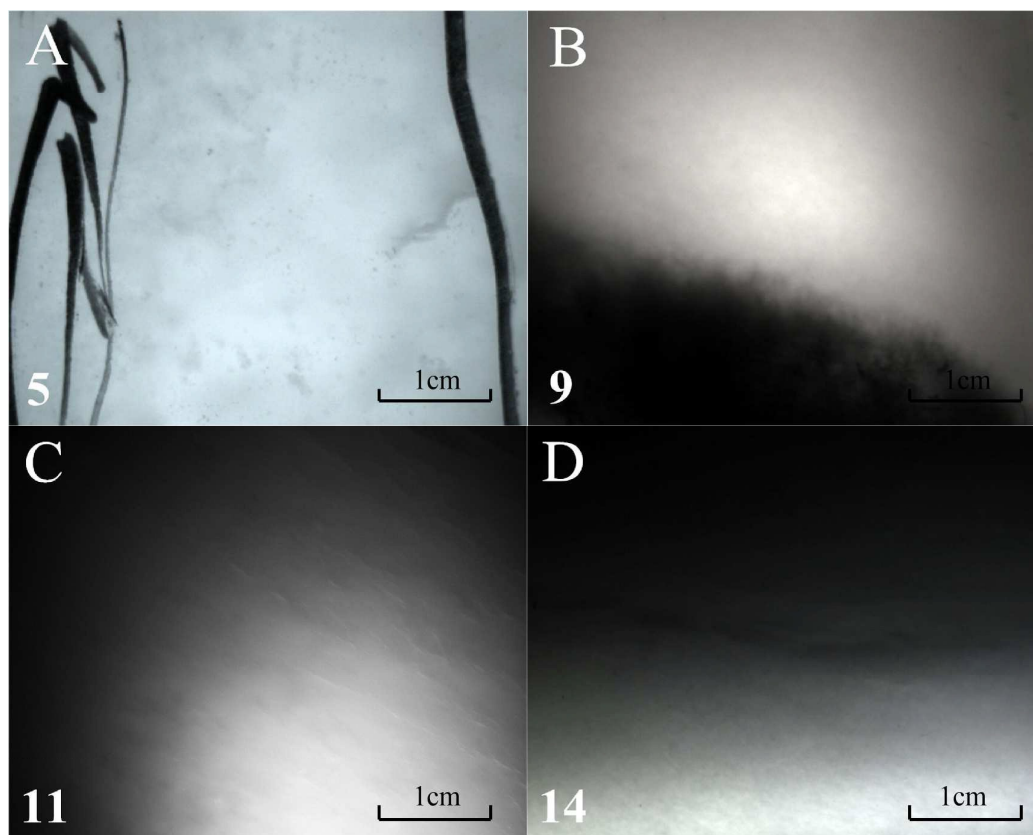
Fig. 7 The selected original gray pictures of nephrite samples from Xinjiang, sample A is from Ruoqiang (the black lines is made with a marker pen), sample B is from Qiemo, sample C is from Yecheng, sample D is from Hetian.

4.2. Clustering results

First of all, we did principal component analysis (PCA) on the data obtained by GLCM algorithm. The first three principal components (PCs) accounted for 47.556%, 26.722% and 19.453% of the contribution rate separately which amount to 93.781%. We just selected the first two PCs to make a scatter plot and got a good classifying quality (Fig. 8).
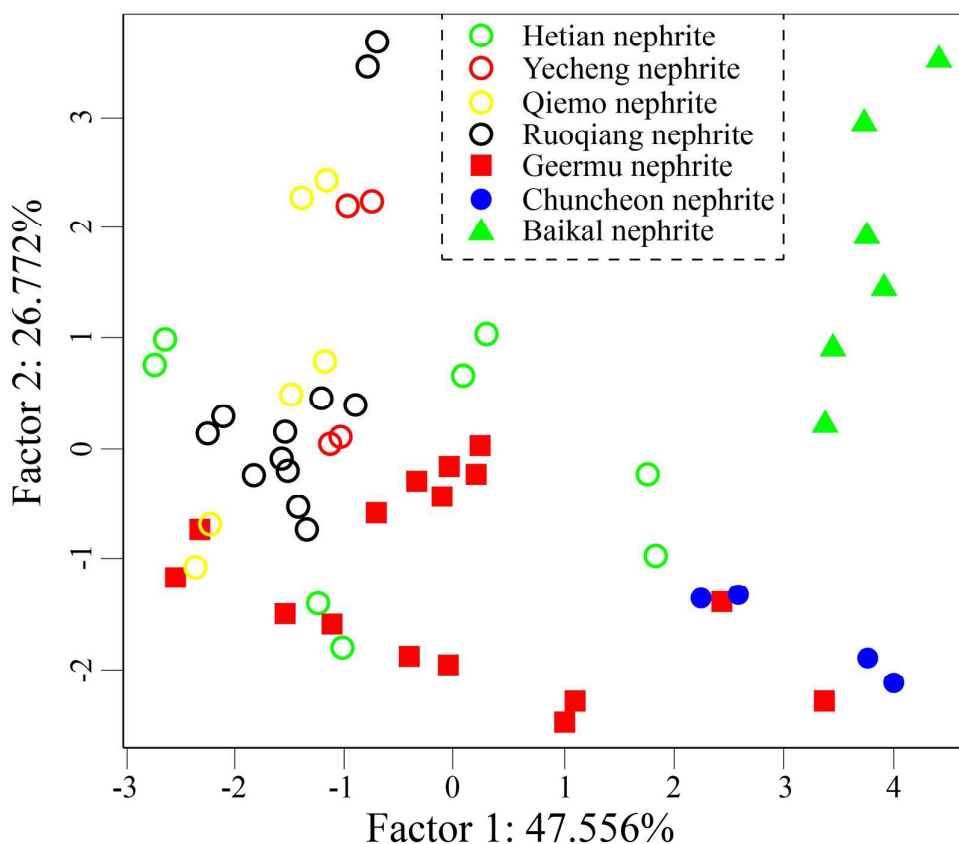
Fig. 8 The two-dimensional scatter plot of principal component analysis of nephrites from different localities.

As we can see in Fig. 8, the Baikal nephrite shows a very high degree of discrimination and there is no overlapping of nephrite from other provenances. The situation in Chuncheon nephrite is similar but a few Qinghai nephrite resemble to them. There is another quite important phenomenon that many scattered points are distributed in pairs. During the measuring process, we selected two regions in each sample for analysis. Hence it indicats that each sample has a homogeneity in GLCM algorithm.

Next, we adopted the hierarchical clustering. In order to display the clustering process of different samples and make full use of data, we made a dendrogram with the Euclidean square distance and the inter group average connection method, as shown in Fig. 9. The longitudinal axis of the dendrogram is the number and provenance of the samples while the horizontal axis represents the relative distance of the various categories which is the result of resetting according to the distance ratio. The relative distance directly quantifies the difference between different samples, and the similar samples will get together in the earliest clustering process.
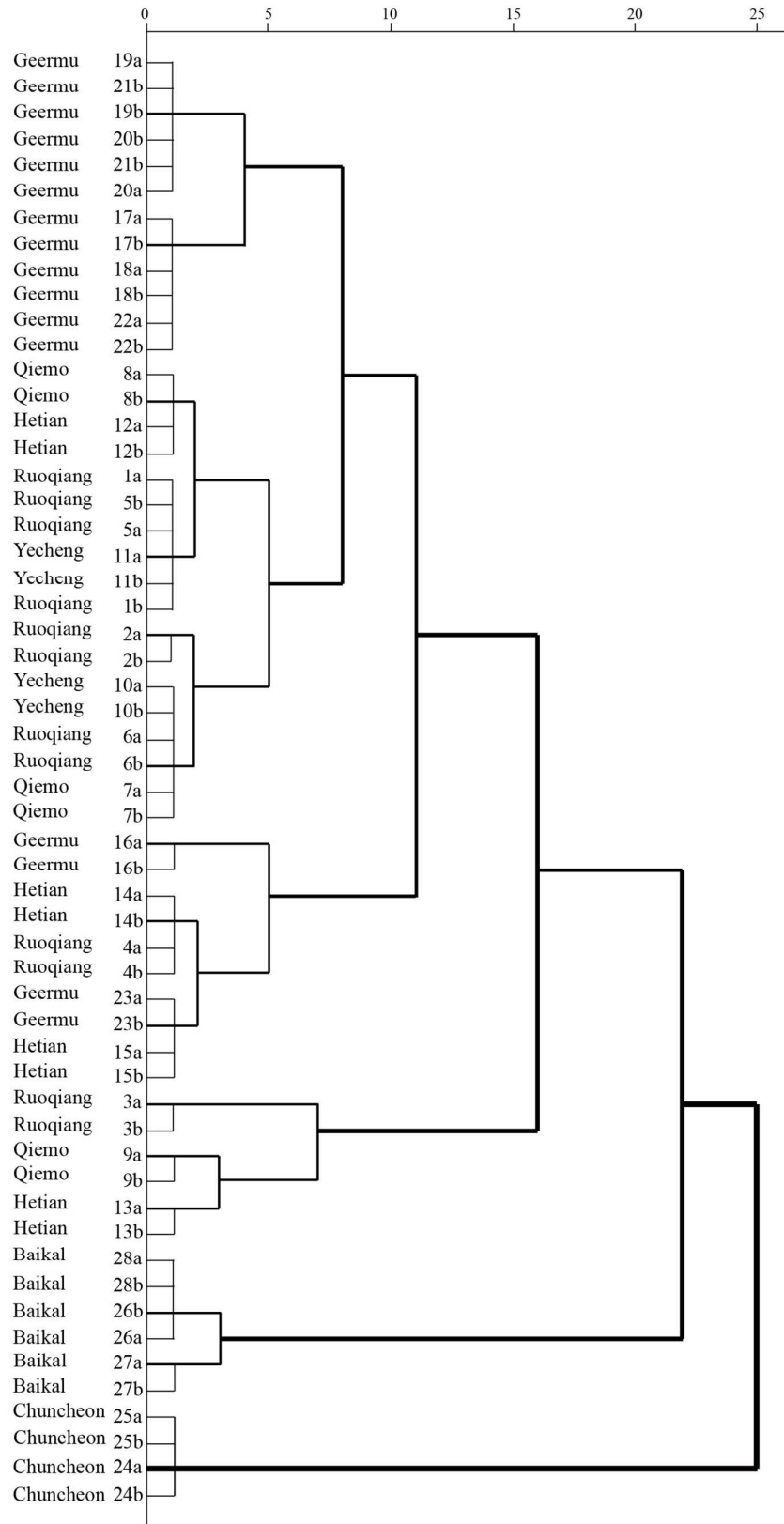
279

280   **Fig. 9** The dendrogram of 28 analyzed nephrite samples.

281 For instance, the distance between nephrite samples from Chuncheon is only 1, so they are
282 divided into a class at the very beginning. But until the last process the small group finally merge
283 with the other big group, here the distance of these two groups already becomes 25 which shows
284 that there are huge and obvious differences between them. Similarly, nephrite samples from Baikal
285 are also grouped together in the initial stage since the distance is just 1. However, they are
286 clustered with other samples at a very later stage only earlier than Chuncheon nephrite. At this
287 time, the distance between nephrite samples from Baikal and nephrite samples from Xinjiang and
288 Qinghai is 22, a little smaller than 25. Thus it can be seen that the characteristics of Chuncheon
289 nephrite are the most unique, which can be distinguished from other different places, and then
290 Baikal nephrite also has considerable differences. This conclusion should be more accurate than
291 than that obtained from the scatter plot in Fig. 8.

292 Moreover, the two data of the same sample always form a pair at the first beginning, in spite
293 of the occasional exceptions such as No. 1a and No. 1b. However, the uniformity of each sample
294 is an objective existence.

295 What remains to be discussed is the nephrite samples from Xinjiang and Qinghai. The
296 distance between the two major groups of the dendrogram except nephrite from Chuncheon and
297 Baikal is 16, much smaller than 22 and 25.. Although the nephrite from Qinghai and Xinjiang can
298 not be completely differentiated, there is also a significant difference between each other.. Because
299 the samples of numbers from 17 to 21 composed a small group indicating they have some general
300 characters. What is not enough is that the discrimination of Xinjiang nephrite in four specific
301 provenances is not very accurate. There is often a small group made up with samples from
302 multiple provenances. Perhaps some respective characteristics exist in Xinjiang nephrite of
303 different specific provenances.

304 Multivariate statistical analysis depends on the selection of characteristic parameters in
305 GLCM algorithm. These parameters chosen in this study maybe not completely cover all the
306 texture features of the image. However, as a very special class of samples, some parameters can
307 not effectively reflect the characteristics of nephrite. And the characteristic parameters are
308 influenced by the step length and the direction, so it is actually a complex parameter system.

309 Moreover, there are many other factors that will affect the cluster analysis such as abnormal
310 data, the collinearity between the variables, the weight of variables, the fluctuation of data and so
311 on. But the homogeneity of nephrite texture guarantees the stationarity of the data from the result.
312 In addition, if some variables are correlated, the cluster analysis will be ineffective or unaffected.
313 Because the correlated variables can be regarded as more weighted. For example, when we do
314 PCA, these correlated variables will constitute a PC or PCs with higher factor. Due to the
315 limitation of the sample size, we have only initially proposed a supervised statistical model. This
316 does not take the full advantages of GLCM. If there is more data, we can use unsupervised ways
317 to find better parameter combinations to avoid the interference caused by the problem of
318 collinearity.

319

Analytical Methods

320 **5. Conclusions**

321   This study shows that multi-spectral imaging technology and gray-level co-occurrence matrix
322 can be effectively applied to non-destructive testing and image structure analysis as a new method
323 for identification of nephrite provenance.

324   Firstly, multi-spectral imaging can reveal the interior submicro-structure image of nephrite,
325 and preserve the reliable empirical data for visual theory which work well in the process of
326 identification. Secondly, the gray-level co-occurrence matrix can give a quantitative description of
327 the texture features in nephrite. Finally, the multivariate statistical analysis by the selected
328 characteristic parameters can effectively complete the clustering discrimination of the samples.

329   This research shows that nephrite from Chuncheon, South Korea and nephrite from Baikal,
330 Russia have strong features respectively, easy to be distinguished from the nephrite of Xinjiang
331 and Qinghai. Besides, the nephrite from Geermu, Qinghai belong to a category of their own and a
332 few nephrite from Xinjiang overlap with them.

333   Nevertheless, the samples in this analysis are very limited, which is far from enough to
334 establish a typical structure image database. Moreover, the image obtained by the current
335 multi-spectral imaging technology needs to be further processed to make the image clearer.
336 Though the statistical model is not necessarily the most reasonable, the clustering discriminant of
337 the sample is effective. We hope to optimize the algorithm and make a good discriminant model in
338 the future.

339

348

349 **References**

350

351 [1] H. Yang, G.H. Liu, Apreliminary discussion on Xinglongwa jades, in The Origin of Jades in East Asia (Jades of
352   the Xinglongwa Culture). Chinese University of Hong Kong Press, Hong Kong, 2017, 210-216.

353 [2] G. Wen, Z.C. Jing, Chinese neolithic jade: a preliminary geoarchaeological study. Geoarchaeology, 1992, **7**
354   251-275.

355 [3] Y. Liu, J. Deng, G.H. Shi., T.F. Yui, G.B. Zhang, M. Abuduwayiti, L.Q. Yang, X. Sun, Geochemistry and
356   petrology of nephrite from Alamas, Xinjiang, NW China. Journal of Asian Earth Sciences, 2011, **42**, 440-451.

357 [4] X.X. Ling, E. Schmädicke, R.H. Wu, S.Q. Wang, J. Gose, Composition and distinction of white nephrite from
358   Asian deposits. Neues Jahrbuch für Mineralogie, 2013, **190**, 49-65.

359 [5] Z.W. Zhang, F.X. Gan, H.S. Cheng, PIXE analysis of nephrite minerals from different deposits. Nuclear
360   Instruments and Methods in Physics Research B, 2011, **269**, 460-465.

361 [6] H. Yu, R. Wang, J. Guo, J. Li, X. Yang, Color–inducing elements and mechanisms in nephrites from Golmud,
362   Qinghai, NW China: Insights from spectroscopic and compositional analyses. Journal of Mineralogical &
363   Petrological Sciences, 2016, **111**, 313-325.

364 [7] T.H. Chen, T. Calligaro, S. Pagès-Camagna, Menu, M., Investigation of Chinese archaic jade by PIXE and
365   µRaman spectrometry. Applied Physics A, 2004, **79**, 177-180.

366 [8] J. Yu, Z. Hou, S. Sahar, J. Dong, W. Han, T.J. Lu, Z. Wang, Provenance classification of nephrite jades using
367   multivariate LIBS; a comparative study. Analytical Methods, 2017, **10**(3).

368 [9] B. Siqin, R. Qian, S. Zhuo, J. Gao, J. Jin, Z.Y. Wen, Studies of rare earth elements to distinguish nephrite
369   samples from different deposits using direct current glow discharge mass spectrometry. Journal of Analytical
370   Atomic Spectrometry, 2014, **29**, 2064-2071.

371 [10] X.X. Ling, E. Schmädicke, Q.L. Li, J. Gose, R.H. Wu, S.Q. Wang, Y. Liu, G.Q. Tang, X.H. Li, Age
372   determination of nephrite by in-situ SIMS U–Pb dating syngenetic titanite: A case study of the nephrit deposit
373   from Luanchuan, Henan, China. Lithos, Volumes 220-223 ,2015, 289-299.

374 [11] S.Q. Wang, L.H. Sun, Visual Identification of Tremolite Features of Five Origins in Today's Nephrite Jade
375   Market. 2013 China Symposium on jewellery academic exchange, Beijing, China, 2013 (in Chinese with
376   English abstract).

377 [12] F.H. Imai, M.R. Rosen, R.S. Berns, Comparative Study of Metrics for Spectral Match Quality. Conference on
378   Colour in Graphics, 2002, **5**, 492-496.

379 [13] C. Fischer, I. Kakoulli, Multispectral and hyperspectral imaging technologies in conservation: current research
380   and potential applications. Studies in Conservation, 2006, **51**, 3-16.

381 [14] G. Themelis, J.S. Yoo, V. Ntziachristos, Multispectral imaging using multiple-bandpass filters. Optics Letters,
382   2008, **33**, 1023-1025.

383 [15] J. Wang, Studies On Multi-Spectral Image Fusion And Application (Masters' Thesis). Wuhan Institute of
384 Technology, Wuhan, Hube, 2005, 4-7.

385 [16] H. Liang, Advances in multispectral and hyperspectral imaging for archaeology and art conservation. Applied
386   Physics A, 2012, **106**, 309-323.

387 [17] A. Pelagotti, D.A. Mastio, D.A Rosa, A. Piva, Multispectral imaging of paintings. Signal Processing Magazine
388   IEEE, 2008, **25**, 27-36.

389 [18] R. Hedjam, M. Cheriet, Historical document image restoration using multispectral imaging system. Pattern
390   Recognition, 2013, **46**, 2297-2312.

391 [19] L.G. Shapiro, G.C. Stockman, Computer Vision. Prentice-Hall, Upper Saddle River, New Jersey, 2001.

392 [20] R.M. Haralick, Statistical and structural approaches to texture. Proeeedings of the IEEE. 2005, **67**, 786-804.

393 [21] L.K. Soh, C. Tsatsoulis, Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices.
394   IEEE Transactions on Geoscience & Remote Sensing, 1999, **37**(2):780-795.

395

396

397

398

399

400

401