Qualitative Interpretation of Spectral Images: Reasoning with Uncertain Evidence

Qi Zhao and Toyoaki Nishida
Graduate School of Information Science
Nara Institute of Science and Technology
8916-5, Takayama-cho, Ikoma-shi, Nara 630-01, Japan
{zhao, nishida}@is.aist-nara.ac.jp

Abstract

Interpreting spectral images requires comparing known patterns with input data (images) to identify which patterns are contained in the input data. In practice, however, it is hard to identify any pattern when the inaccuracy of input data is not slight. In this paper, we present a method for interpreting spectral images by using qualitative reasoning. First, we put forward a new concept called support coefficient function *(SCF)* which can be used to extract, represent, and calculate qualitative correlations among data. Then, we introduce an approach to determining dynamic shift intervals of inaccurate data on the basis of qualitative correlations. Finally, we discuss how to use qualitative correlations as evidence of enhancing or depressing hypotheses for inaccurate data. The method has been applied to a practical system for interpreting infrared spectral images. We have fully tested the system against several hundred real spectral images. The rate of identification *(RI)* and the rate of correctness *(RC)* are near 90% and 74% respectively, and the latter is the highest among known systems.

1 Introduction

Interpreting spectral images is a special problem of diagnosis. The problem requires comparing known patterns with input spectral images to identify which patterns may be contained by the images [Anand *et al.*, 1991; Sadtler, 1988]. Because spectral data are always inaccurate, one difficult task is to deal with uncertain evidence.

Currently known methods and systems of spectral image interpretation are primarily based on quantitative analysis [Culthup *et al.*, 1990; Sadtler, 1988]. The essential principle of quantitative analysis is to determine the possibility that a pattern may be contained by an image by calculating the difference between the pattern and the parts of the image. In practice, however, a critical problem of applying quantitative analysis is that it is hard to identify any pattern when the inaccuracy of input data is not slight. Fuzzy logic and probability theory can partially solve the problem [Duda and Nilsson, 1976; Zadeh, 1978], but they can not consider qualitative correlations among data[1] which are very important and effective in spectral image interpretation.

We present a novel method for interpreting spectral images by using qualitative reasoning. The method draws inferences on the basis of qualitative features of spectral images, and uses qualitative correlations among data as evidence when input data are inaccurate.

We put forward a new concept called support coefficient function *(SCF)*. *SCF* can be used to extract, represent, and calculate qualitative correlations among data. On the basis of qualitative correlations and dynamically obtained information, shift intervals of inaccurate data can be dynamically determined. When input data are inaccurate, qualitative correlations can provide evidence to enhance or depress hypotheses for inaccurate data [Zhao and Nishida, 1994],

We have developed a practical system on infrared spectral image interpretation by using the method, and have fully tested the system against several hundred real spectral images. The experimental results with the system are excellent. Both the rate of correctness *(RC)* and the rate of identification *(RI)* increase significantly by using the method[2].

In the following sections, we first describe the problem of spectral image interpretation in section 2. Then in section 3, we address the essentials of our method including some useful definitions. In section 4, we introduce our method for interpreting inaccurate spectral data by considering qualitative correlations among data. Section 5 demonstrates the application of the method to a system on infrared spectral image interpretation, and shows the experimental results with the system. Related work is discussed in section 6. Our ongoing research is briefly introduced in section 7. Conclusions are addressed in section 8.

2 Motivating Problem

The primary task of spectral image interpretation is to identify what patterns a spectral image contains by inter-

[1] Detailed definition will be given in section 3.1.

[2] *RC* and *RI* are two important metrics for evaluating the solutions of infrared spectral image interpretation. We will give the detailed definitions of *RC* and *RI,* and show the experimental results of our system in section 5.2.

preting the spectral image. In the rest of the paper, we limit the problem to interpreting infrared spectral images of unknown compounds to identify what partial components *(PC)* the unknown compounds contain without loss of generality.

Briefly, the infrared spectral image of an unknown compound can be thresholded and represented as a set of peaks:

$$Sp = \{p_1, p_2, ..., p_n\},$$

where every peak consists of its frequency position (*f*), strength (*s*), and width *(w)* respectively:

$$p_i = (f_i, s_i, w_i), \quad i = 1, 2, ..., n.$$

The patterns of known partial components can be represented in the following form:

$$PC_j = \{p_{j_1}, p_{j_2}, ..., p_{j_m}\} = \{(f_{j_l}, s_{j_l}, w_{j_l}) \mid l = 1, 2, ..., m\},$$

which means if PC_j is contained by a compound, then on the spectral image of the compound, there will be peak $p_{j_1}, p_{j_2}, ...,$ and p_{j_m}

Ideally, if all peaks on spectral images are accurate, the process of interpreting infrared spectral images of unknown compounds can be simply described in the following way.

1. Select a peak from *Sp.* Retrieve all partial components whose patterns have the same peak, and put the partial components in a candidate list: *CL;*
2. Select a partial component, PC_j, from *CL.* If $PC_j \subset Sp$, then put the partial components in a solution list: *SL;* otherwise, delete the partial component from *CL;*
3. Goto 2 until all partial components in *CL* are checked;
4. Goto 1 until all peaks on *Sp* are identified;
5. Delete conflicts (overlaps) among partial components in *SL,* and output *SL* as the solution[3].

The overview of the process is shown in Figure 1.

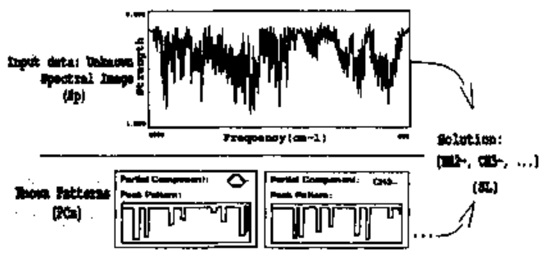

Figure 1: Overview of interpretation process

In practice, however, spectral data are often inaccurate due to various reasons most of which are unforeseen, or unknown at all. Therefore, the peaks created by a partial component on real spectral images are often different from the peaks described by the pattern of the partial component.

Fuzzy logic provides mathematical fundamentals of representation and calculation of inaccurate data [Bowen *et al.*, 1992; Negoita and Ralescu, 1987; Zadeh, 1978]. For example, peaks in the patterns of partial components may be described in a fuzzy fashion like f_a *is around* f_o (not f_a *is equal* f_o), and a fuzzy region may also be defined to represent the peaks like $[f_o\text{-}\Delta f,\ f_o+\Delta f]$. Probabilistic reasoning provides a practical framework for reasoning under uncertainty and reasoning with inaccurate data [Dempster, 1968; Pearl, 1987]. In many systems, subjective statements are used to take the place of statistics of data or evidence when statistical samples are insufficient or absent, such as certainty factors in MYCIN [Shortliffe, 1976] and prior probabilities in PROSPECT [Duda and Nilsson, 1976].

However, both fuzzy logic and probability theory are based on quantitative features, but in many cases, especially when data are seriously inaccurate, qualitative features are much more important than quantitative ones. Figure 2 shows two simple cases. The quantitative difference between the two spectral images in (a) is smaller than that between the two spectral images in (b). By conventional fuzzy or probabilistic methods, the two images in (a) should be closer than those in (b). Actually, the two spectral images in (b) may be created by the same partial component in some cases, while the two spectral images in (a) are definitely created by two different partial components because the number of peaks is different. Here, the number of peaks on spectral images is a critical qualitative feature.

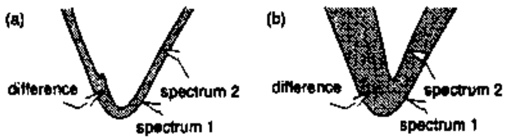

Figure 2: Qualitative features of spectral images

We propose a novel method to interpret infrared spectral images by using qualitative reasoning. The method draws inferences on the basis of qualitative features of spectral images, and uses qualitative correlations among data as confirmatory or disconfirmatory evidence when input data are inaccurate.

Suppose *P(a)* represents the possibility that two images in (a) are created by the same partial component, and *P(b)* represents the possibility that two images in (b) are created by the same partial component. At the beginning, *P(a)* > *P(b),* but with the obtaining of qualitative correlations among data, *P(b)* will increase, and *P(a)* will decrease.

By our method, the above ideal process can be briefly expressed as the following predicate calculus:

$$\forall p_i \forall PC_j((p_i @ PC_j) \wedge (PC_j @ Sp) \rightarrow (PC_j \in SL)),$$

where $p_i @ PC_j$ and $PC_j @ Sp$ are two essential qualita-

[3] We will not discuss conflict-resolving in this paper, but concentrate on the method for identifying inaccurate data.

tive predicates in our method which represent that p_i qualitatively belongs to PC_j (i.e., ? $p_i \in PC_j$), and PC_j qualitatively belongs to Sp (i.e., ? $PC_j \subset Sp$), respectively. Realizing $A@B$ is based on qualitative correlations among data and dynamic information obtained in interpretation.

3 Preliminaries

3.1 Qualitative Correlations among Data

Definition 3.1 *Related data:* If data d_1, d_2, ..., and d_m describe a common phenomenon altogether, or they refer to the same behavior simultaneously, then they can be treated as related data.

In infrared spectral image interpretation, there are two types of related data: (1) as far as a single peak is concerned, the frequency /, strength s, and width *w* of the peak are related data; (2) if the pattern of a partial component is considered, all the peaks in the pattern are related data.

Definition 3.2 *Qualitative correlations among related data:* If d_j and *dj* are two related data, then the presence of *d{* somewhat (qualitatively) enhances the presence of d_j, and the absence of d_i somewhat (qualitatively) depresses the presence of *dj.* The above effects are called qualitative correlations among related data.

In infrared spectral image interpretation, if data gathered from spectral images look like, but are not exactly the same as, the data described in the patterns of partial components, considering qualitative correlations among related data can obtain qualitative evidence. For example, suppose the frequency position of a peak is different from that in a pattern, but both the strength and width of the peak are the same as described in the pattern, then the frequency position of the peak may still be identified because both of its related data support it. Similarly, if peaks at low frequency sections are inaccurate, considering their related peaks at high frequency sections may help identify these peaks, and vice versa.

3.2 Support Coefficient Function

Definition 3.3 *Support coefficient junction (SCF):* If there are m - 1 data related to d_i, then the support coefficient function of d_i calculates the total effects from the related data by considering the qualitative correlations between d_i and each of its related data.

Suppose $\pi(d_i, d_j)$ represents the qualitative correlation between d_i and d_j, then the support coefficient function of d_i can be defined as:

$$SCF_i = \beta(\sum_{j=1, j\neq i}^{m} \pi(d_i, d_j), m).$$

SCF_i should directly depend on how many and how much related data support d_t. When SCF_i is greater than a certain value given by domain experts, the related data tend to support d_i otherwise, the related data tend to depress *di.*

3.3 Evidence Based on SCF

In section 2, we used $p_i@PC_j$ to express that *pi* can be qualitatively identified from PC_j. Realizing $p_i@PC_j$ requires to define a shift interval for PC_j like:

$$PC_j \pm \Delta = \{(p_{j_l} \pm \Delta) \mid l = 1, 2, ..., m\},$$

and then to determine the possibility of $p_i \in PC_j \pm \Delta$.

The above formula is similar to that in fuzzy logic, but contains completely different meanings. The primary difference is that the shift intervals are dynamically determined by *SCFi.*, while in fuzzy logic, the fuzzy regions are usually provided by domain experts in advance.

Definition 3.4 *Shift interval:* Shift interval is a dynamic region for inaccurate data. Given a standard fuzzy region for inaccurate di, the shift interval of *di* varies around the standard fuzzy region on the basis of *SCFi.* When *SCFi* shows that the related data support *di* the shift interval of *di* becomes wider than the standard fuzzy region. On the other hand, when *SCFi* shows that the related data do not support di, the shift interval of *di* becomes narrower than the standard fuzzy region.

Definition 3.5 *Evidence based on SCF: SCFi* determines the shift interval of di, that is, *SCFi* determines how widely d_i is allowed to shift. The wider the shift interval, the more easily d_i is identified. Therefore, *SCFi* provides confirmatory or disconfirmatory evidence for identifying d_j.

4 Qualitatively Interpreting Spectral Images

As described in section 3, there are two types of related data in infrared spectral image interpretation. First, all the features of a peak are related data. Second, all the peaks created by the same partial component are also related data. For making our introduction brief, in this section, we omit the process of considering the former type of related data, and directly introduce the process of considering the latter one.

Suppose *m* peaks (p_1, p_2, ..., P_m) may be created by the same partial component, and d_0 is the standard fuzzy region for real peaks, then our method can be described as the following steps.

[Step 1] Defining *support coefficient function*

First, we define the qualitative correlation between two related peaks, p_i and p_j, as:

$$c_i(p_j) = \begin{cases} 1 & \text{if } p_j \text{ can be found in} \\ & \quad [p_j - d_o, p_j + d_o] \text{ on } Sp \\ 0 & \text{if } p_j \text{ can not be found in} \\ & \quad [p_j - d_o, p_j + d_o] \text{ on } Sp, \end{cases}$$

where $C_i(p_j)$ expresses the qualitative effect of p_j on p_i'. *Ci(pj)=l* means that *pi* is enhanced since its related data item *pj* can be identified; $c_i(p_j)$=0 means that p_i is depresses since its related data item *pj* can not be identified. *Ci(pj)* provides dynamic information for making interpretation. The definition of *Ci(pj)* is simply based on the consideration that if a peak of a partial component has been identified, then the peak will support the coexisting peaks (related peaks) that the partial component may create at the same time.

Based on *Ci(pj)* (j = 1, 2, ..., m, $j \neq i$)„ we define the support coefficient function of *pi* in the following way.

$$SCF_i = \frac{1 + \sum_{j=1, j \neq i}^{m} c_i(p_j)}{m},$$

where $0 < SCF_i \leq 1$, and SCF_i expresses the qualitative correlations between p_i and all of its related peaks.

If $m = 1$, then $SCF_i = 1$. When $m > 1$, SCF_i is in the direct ratio to the number of related peaks which may be identified.

[Step 2] Defining *Shift interval*

$$\Delta d_i = \frac{(2m-1)d_o}{m} \times SCF_i,$$

where $0 < \Delta d_i < 2d_o$, and Δd_i is in the direct ratio to SCF_i.

If $m = 1$, then $SCF_i = 1$, and $\Delta d_i = d_o$. When m is fixed, the more the related peaks are identified, the greater SCF_i is, therefore the greater Δd_i is. When SCF_i is fixed, Δd_i depends on the number of related peaks.

[Step 3] Making *qualitative hypotheses*

On the basis of SCF_i and Δd_i, we can make qualitative interpretation for inaccurate data. A triple is used to represent the interpretation:

$$IN = \{(p_i, p_{j_i}, \mu_i) \mid i = 1, 2, ..., m\},$$

where p_{j_i} is a peak described in the pattern of a partial component (or in other words, p_{j_i} is a peak that the partial component should create); p_i is a real peak on the input spectral image which is suspected to be p_{j_i}; and μ_i is the possibility of p_i being accepted as p_{j_i} (i.e., $p_i@PC_j$) which is defined as:

$$\mu_i = max\left\{0, 1 - \frac{|p_i - p_{j_i}|}{\Delta d_i}\right\},$$

where $0 \leq \mu_i \leq 1$.

The following properties can be drawn from the above formulas. These properties clearly explain how qualitative correlations among related data provide evidence of making interpretation for inaccurate data.

Property 4.1: With the same m, the more the related peaks are identified, the greater SCF_i is; otherwise, the smaller SCF_i is.

Property 4.2: With the same m, the greater the SCF_i, the greater is Δd_i, which implies that the more the related peaks support p_i, the more widely p_i is allowed to shift[4].

Property 4.3: With the same p_i, the greater the Δd_i, the greater is μ_i, which implies that the wider the dynamic shift interval, the greater is the possibility of p_i being identified as p_{j_i}. Formally, if $\Delta d_i'' \geq \Delta d_i' \geq \Delta d_i$, then $\mu_i'' \geq \mu_i' \geq \mu_i$.

Property 4.4: (Based on Property 4.2 and Property 4.3) SCF_i provides evidence for accepting or rejecting inaccurate p_i because μ_i is in the direct ratio to Δd_i, and Δd_i is in the direct ratio to SCF_i.

Actually, IN represents a qualitative hypothesis that partial component PC_j ($PC_j = \{p_{j_1}, p_{j_2}, ..., p_{j_m}\}$) is contained by the input image (i.e., $PC_j@Sp$). The possibility of the hypothesis can be easily calculated from μ_1, μ_2, ..., and μ_m:

$$P(PC_j) = \frac{\sum_{i=1}^{m} s_i \times \mu_i}{\sum_{i=1}^{m} s_i}, \quad s_i \geq 0,$$

where constant s_1, s_2, ..., and s_m represent the priorities of the related data. For example, the peaks of CH_3 in frequency section of 2800 cm^{-1} to 3000 cm^{-1} are more distinct than its peaks elsewhere, as a result, the peaks in this section are prior to those located in other sections.

5 A System on Spectral Image Interpretation

We have developed a practical system on infrared spectral image interpretation by applying the method, and have fully tested the system against several hundred real spectral images. The experimental results with the system are excellent.

5.1 Overview of the System

The system is implemented with C and MS-WINDOWS. Figure 3 shows the data flow diagram of the system. The input data of the system are gathered from given spectral images, and the solutions are partial components that the input images may contain.

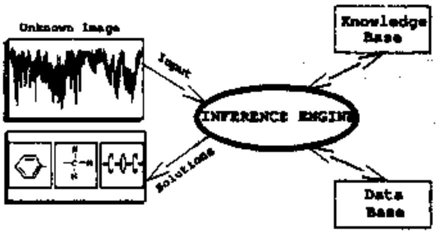

Figure 3: Data flow diagram of the system

Because inferences are based on qualitative features and qualitative correlations among related data, the system can obtain very high correct interpretation performance with noisy images or images of mixed compounds.

5.2 Experimental Results

We have compared two methods in the experiments. The first method (called "AF") is a conventional fuzzy method which is used by most similar systems [Clerc *et al.*, 1986]. *AF* adopts empirical fuzzy regions, and does not use *SCF*. The second method (called "*AF**") is the

[4]Related properties can be drawn as: **Property 4.2a**: With the same SCF_i, the greater the m, the less Δd_i varies along with m, which implies that the greater the number of related peaks, the less a single peak can affect p_i. **Property 4.2b**: Δd_i is in linear relation to SCF_i with a slope that is equal to, or greater than 1.5, which implies that Δd_i heavily depends on SCF_i. **Property 4.2c**: Along with the increase of m, the slope increases very slightly, which implies that Δd_i depends on the number of the related peaks which support p_i, rather than the total number of related peaks.

method proposed in this paper which uses *SCF* to provide qualitative evidence for identifying inaccurate data.

We have tested *AF* and *AF** against several hundred real infrared spectral images of organic compounds. The experimental results show that *AF** is significantly better than *AF.*

Table 1 lists part of the experimental results in which the first column shows the correct solutions; the second column indicates the solutions obtained by *AF]* and the third column shows the solutions obtained by *AF*.*

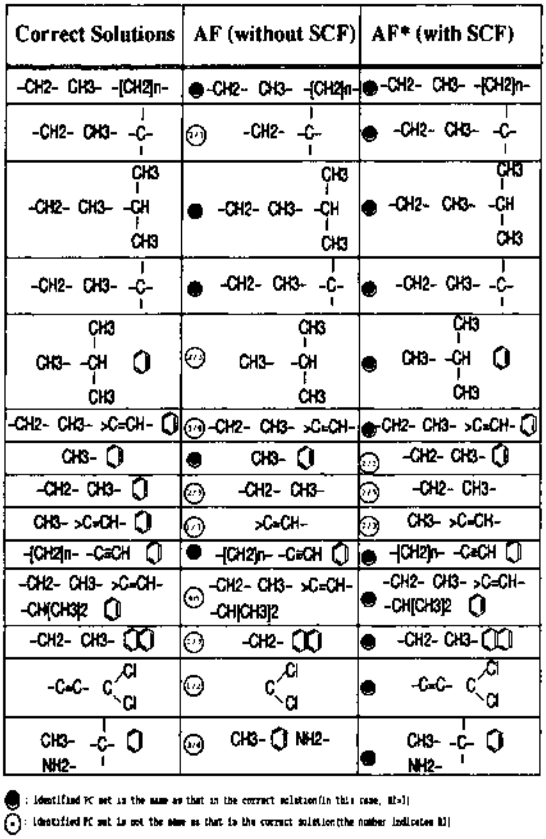

Table 1: Experimental results with *AF & AF**

There are two important standard metrics for evaluating the solutions of infrared spectral image interpretation for unknown compounds:

Definition 5.1 *Rate of correctness (RC):* the rate that the identified partial component set is exactly the same as the partial component set in the correct solutions.

Definition 5.2 *Rate of identification (RI):* the rate that the partial components in the correct solutions are identified.

Table 2 shows the comparison between *AF* and *AF** with the two standard metrics. The comparison demonstrates that both the *RC* and the *RI* increase by integrating *SCF,* and the *RC* increases more significantly. The reason is that although *AF* can identify most patterns on spectral images, the rate that it can identify all patterns on spectral images is low because in practice there are always some partial components whose real peaks seriously shift from the expected values.

	RC	(error-rate)	*RI*	(error-rate)
AF	0.455	(0.545)	0.812	(0.188)
*AF**	0.736	(0.264)	0.894	(0.106)

Table 2: Evaluation of *AF* & *AF with *RC* and *RI***

So far as we know, the *RC* of our system is the highest among the similar systems, and the *RI* of our system is higher than that of many similar systems [Clerc *et* a/., 1986; Hasenoehrl *et* a/., 1992; Puskar and Levine, 1986; Sadtler, 1988],

5.3 An Example

It is interesting to discuss the following example shown in Figure 4.

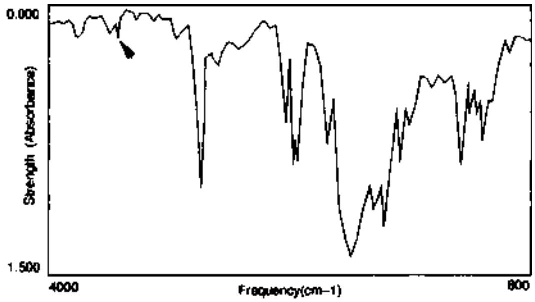

Figure 4: An example of spectral image interpretation

Figure 4 shows a spectral image of an unknown compound (Anisole). The peak with an arrow is created by partial component CH_3. However, the peak can not be identified by conventional methods because the peak of CH_3 at this frequency position should be a strong peak (i.e., s > 1.000), but in this example, the strength of this peak is very weak (5 < 0.130). So the membership degree of this peak is very close to zero, if not equal to. By using our method, the peak is identified as the peak of CH3 because the support coefficient of this peak is very high *(SCF* = 0.93).

6 Comparison with Related Work

Traditional methods and systems of spectral image interpretation are based on numerical analysis which identify partial components by calculating the quantitative similarity or closeness between the patterns of known partial components and the input spectral images [Clerc *et* a/., 1986; Culthup *et* a/., 1990; Sadtler, 1988]. As we have discussed in section 2, in most cases especially when the inaccuracy of data is not slight, qualitative features of spectral images are much more important and effective than quantitative ones. In other words, the similarity/difference between two spectral images does not strictly reflect the similarity /difference between two structures. Therefore, in general, the solutions of the systems based on numerical analysis are only a series of candidates from which users have to finally decide the possible ones by themselves.

Many recently known methods and systems are based on AI techniques [Anand *et al,* 1991; Hasenoehrl *et al,* 1992; Puskar and Levines, 1986]. Common techniques mainly include production systems, fuzzy logic and neural networks. Although various reasoning and interpreting methods have been studied, the approaches to dealing with inaccurate data in this kind of methods and systems are almost the same, i.e., reference values and fuzzy regions of inaccurate data are empirically provided in advance. Thus, qualitative correlations among data and dynamic information can not be used properly. As a result, compared with our system, the limitation of this kind of methods and systems is that, generally, they are only effective for a class of compounds, or pure compounds because when image data are seriously inaccurate, many useful inferences can not be drawn.

7 Ongoing Research

The reasons causing image data inaccurate are various. However, the interaction among partial components themselves is an important factor or, in other words, the influence from coexisting partial components may be a chief reason to cause the peaks of a partial component to shift from their theoretical values. In practice, spectroscopists frequently use the knowledge like: *"if C_6H_6 coexists with CH_3, then the peaks of CH_3 around 2900 cm^{-1} may shift."*, or *"if -C-O-C- has been identified, then the pattern of CH_3 may change."* Therefore, it is possible to update the possibilities of identified partial components by considering the interaction among them after qualitative interpretation has been made.

Take the spectral image shown in Figure 4. The image contains partial component C_6H_5- (benzene-ring), *-C-O-C-*, CH_3 and others. On the spectral image, the peak of CH_3 around 2900 cm^{-1} is seriously inaccurate due to the influence of C_6H_5- and *-C-0-C-*. Although CH_3 is identified based on qualitative correlations among related data, the possibility of CH_3 being contained by the image is not high $\{P(CH_3) = 55\%)$.

By Bayes rules associated with subjective statements, we can get $P(CH_3 \mid C_6H_5-)$ and $P\{CH_3 \mid -C-0-C-)$ respectively. And finally, $P(CH_3)$ may be updated with $P(CH_3 \mid 5)$, where 5 stands for all relevant observations.

Theoretical and experimental work on the issue is in progress.

8 Conclusions

We have presented a novel method for interpreting spectral images on the basis of qualitative features of spectral images and qualitative correlations among related data. We first put forward a new concept called support coefficient function *(SCF).* Then we proposed an approach to determining dynamic shift intervals, and an approach to calculating possibility of identifying inaccurate data, respectively. A triple, $(p_i, p_j., u_i)$, is used to represent the hypothesis that p_i can be accepted as p_j based on qualitative evidence provided by *SCF.*

We have also introduced a knowledge-based system on infrared spectral image interpretation which is developed by applying the proposed method. The system can successfully identify partial components that unknown images may contain. Several hundred real spectral images have been identified, and the results of implementation are quite encouraging (the *RI* and the *RC* are about 0.90 and 0.74 respectively).

References

[Anand *et al,* 199l] Anand, R., Mehrotra, K., Mohan, C. K., & Ranka, S. Analyzing Images Containing Multiple Sparse Patterns with Neural Networks. *Proc. of IJCAI-91,* pp. 838-843, 1991.

[Bowen *et al,* 1992] Bowen, J., Lai, R., *k* Bahler, D. Lexical Imprecision in Fuzzy Constraint Networks. *Proc. of AAAI-92,* pp. 616-621, 1992.

[Clerc *et al,* 1986] Clerc, J. T., Pretsch, E., & Zurcher, M. Performance Analysis of Infrared Library Search Systems. *Acta/Wien],* II, pp. 217-242, 1986.

[Colthup *et al,* 1990] Colthup, L., Daly, H., & Wiberley, S. E. Introduction to Infrared and Raman Spectroscopy. *Academic Press INC.,* 1990.

[Dempster, 1968] Dempster, A. P. A Generalization of Bayesian Inference. *Journal of the Royal Statistical Society,* Vol.B-30, pp. 205-247, 1968.

[Duda and Nilsson, 1976] Duda, R. O. and Nilsson, N. J. Subjective Bayesian Methods for Rule-Based Inference Systems. *Tech. Note 124, SRI Int.,* 1976.

[Hasenoehrl *et al,* 1992] Hasenoehrl, E. J., Perkins, J. H., & Griffiths, P. R. Expert System Based on Principal Components Analysis for the Identification of Molecular Structures from Vapor-Phase Infrared Spectra. *J. of Anal. Chem.,* Vol.64, pp. 656-663, 1992.

[Negoita and Ralescu, 1987] Negoita,C.V. and Ralescu, D. Simulation, Knowledge-Based Computing, and Fuzzy Statistics. *Van Nostrand Reinhold Company Inc.,* 1987.

[Pearl, 1988] Pearl, J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. *Morgan Kaufmann Publishers,* 1988.

[Puskar and Levine, 1986] Puskar, M. A. and Levine, S. P. Computerized Infrared Spectral Identification of Compounds Frequently Found at Hazardous Waste Sites. *J. of Anal.Chem.,* Vol.58, pp.1156-1162, 1986

[Sadtler, 1988] Sadtler Research Laboratories. *Sadtler PC Spectral Search Libraries,* Product Introduction &: User's Manual, 1988.

[Shortliffe, 1976] Shortliffe, E. H. Computer-Based Medical Consultations: MYCIN. *American Elsevier Publishing Inc.,* 1976.

[Zadeh, 1978] Zadeh, L. A. Fuzzy Set as a Basis for a Theory of Possibility. *Fuzzy Sets Syst.,* Vol. 1, pp. 3-28, 1978.

[Zhao and Nishida, 1994] Zhao, Q. and Nishida, T. A Knowledge Model for Infrared Spectrum Processing. *Proc. ofISITA,* pp. 781-786, 1994.