FORENSIC DATABASE SYSTEM DEVELOPMENT FOR BUILDING PAINTS

UMI KALTHOM AHMAD [1,*], WONG JUN WEI [2]
& ROLIANA IBRAHIM [3]

Abstract. Paint flakes are often found at crime scenes as trace physical evidence that offer significance importance for forensic investigations. Matching of unknown paint flakes with known case samples may provide clues in solving crime cases. However, manual paint sample matching of case and control samples is often slow and inefficient. A paint database system for fast data retrieval is much sought after by forensic scientists. In this study, sixty building paints were examined in terms of their color appearance, solubility and Fourier-transform infrared (FTIR) analysis. Ten diagnostic functional groups were selected in IR analysis in order to discriminate between the paints analyzed. The developed database incorporated data of the paint, paint color, solubility testing and IR diagnostic peaks. These attributes were utilized for data searching, retrieval and matching with the unknown paint samples analyzed. An interactive user interface was designed based on the type of data stored in the database. The system flow follows the sequence of paint color matching, then solubility result comparison, and finally IR diagnostic group matching. The developed paint database system supports easy matching of unknown building paint fragments with that stored in the database.

Keywords: Paint flakes; FTIR; database system; list filtering; matching

Abstract. Serpihan cat kerap dijumpai di tempat kejadian sebagai bukti fizikal surih yang amat penting untuk penyiasatan forensik. Perbandingan antara serpihan cat dari tempat kejadian, dengan sampel kes yang maklumatnya diketahui, dapat mengenal pasti asalannya dan dapat menyediakan petunjuk kepada penyiasat forensik dalam menyelesaikan kes jenayah. Walau bagaimanapun, perbandingan sampel cat dengan cara manual adalah lambat dan tidak cekap. Sistem pangkalan data bagi cat rumah adalah amat diperlukan oleh saintis forensik. Dalam kajian ini, enam puluh cat bangunan telah diperiksa dari segi penampilan warna, kelarutan, dan analisis Fourier-transformasi infra merah. Sepuluh kumpulan berfungsi diagnostik telah dipilih dalam analisis untuk membezakan antara cat yang dianalisis. Pembangunan pangkalan data telah pun menggabungkan data yang berkaitan dengan cat, warna cat, ujian keterlarutan dan puncak

---

[1&2] Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor Darul Ta'azim, Malaysia

[3] Department of Information System, Faculty of Computer Science & Information System, Universiti Teknologi Malaysia, 81310, UTM Johor Bahru, Johor Darul Ta'azim, Malaysia

* Corresponding author: umi@kimia.fs.utm.my

diagnostik **IR**. Maklumat-maklumat ini digunakan untuk pencarian data dan perbandingan dengan sampel yang tak diketahui sumbernya. Antara muka pengguna yang interaktif telah dibangunkan berdasarkan jenis data yang disimpan dalam pangkalan data. Aliran sistem bermula dengan perbandingan warna cat, diikuti dengan keputusan keterlarutan, dan akhirnya perbandingan kumpulan puncak diagnostik **IR**. Sistem pangkalan data yang telah dibangun memudahkan proses perbandingan serpihan cat rumah dari tempat kejadian dengan yang disimpan dalam pangkalan data.

Kata kunci: Serpihan cat; FTIR; sistem pangkalan data; senarai penurasan; padanan

1.0   INTRODUCTION

Database management system like a file system, is a way to manage data in a reliable and efficient manner. Database system greatly reduces data redundancy in order to maintain consistency and increase in speed of data searching. The programming language that supports database system, structured query language (SQL), provides multiple functions in managing dataset stored. Inserting new data, updating existing information, deleting the unused and viewing the data accordingly, are some of the basic features that should be included during the development of database management system.

The need of forensic applications leads to emerging of different types of database, mainly the purpose is for data retrieval. Examples are biological DNA database for criminal, fingerprint database system for individual identification, automobile paint and shoe prints database system. Besides data retrieval, those databases are mainly used to match unknown samples with the ones stored in the databases. Normally the exchange of information and analysis results between forensic communities happens all the time [1]. Sharing of information enable extension of the scope of particular dataset, for not only accessing local data, cross country information is also freely available, practically important during the suspect identification in solving crime cases.

For paint database system, Forensic Laboratory Services (FLS) has developed an automotive Paint Data Query (PDQ) database under The Royal Canadian Mounted Police (RCMP) [2]. This database consists of all the chemical and physical information of original automobile paints. It accumulates information provided by automobile companies and also samples vehicles submitted by other forensic laboratories. Other than physical and chemical data, information like the make, model, year, and assembly part are also stored. Once the police or forensic

investigator obtain any automobile paint sample, they will run through some physical examination and chemical testing. The results obtained will be compared with the dataset in the database. Once there is a match for the unknown sample, it is possible to identify the make, year and model of the vehicle. This will narrow down the search of the suspect's car and shorten the investigation time.

However, household paint database has not yet been developed for the forensic investigation purposes, although there is a need in it. Paint flakes or chip are often been collected by the crime scene investigators. The origin of paint flakes is the main aim for forensic investigator to look into [3]. In order to determine the identity of the unknown samples, comparison and matching with the known paint samples is considers one of the effective methods. However, process of identification is tedious since most of the paint information on hand is in documentation form. Manual sample matching is inefficient and time consuming.

Without an organized paint database, the identification and matching of unknown paint fragments become complicated: physical appearance and color comparison, chemical test result matching and infra red (IR) result pairing. All of the mentioned step can be computerized for automation result comparison. Efficiency of the search result can be increased with the development of household paint database system. Furthermore, a proper database management system eases in storing and retrieving paint information, especially when data size is large. Maintaining data consistency using database is much easier compared to data that exist in documentation form [4]. The main purpose of building the database system is to match the unknown case sample, with that stored in the database.

This study describes the development of a building paint database utilizing a list filtering method.

2.0 EXPERIMENTAL

2.1 Chemicals and Materials

A total of seven organic solvents were used for the solubility test. Chloroform, 1,2-dichlorobenzene, pyridine and acetone were laboratory grade while dichloromethane, dimethylformamide, and methanol were analytical grade from a range of suppliers. Each sample was tested by these seven types of organic solvent.

Table 1 shows the physical properties of the organic solvents used for this chemical test. Building paints samples (sixty samples) of various brands and types were kindly donated by individuals from UTM, Seamaster Paint and Colourland Paint Company in Johor.

2.2 Equipment & Apparatus

Paint flakes were subjected to Thermo Nicolet FT-IR iS10 USA with ATR Sampling for the detection of diagnostic functional groups. The instrument was equipped with a diamond crystal and the spectra were recorded in the spectral range between 4000 cm^{-1} and 650 cm^{-1}. An OMNIC spectra software was utilized for data processing.

Table 1　Physical properties of selected solvents

Solvent	Miscibility with water	Boiling point (°C)	Chemical formula	Polarity Index
Dimethylformamide	Miscible	153	C_3H_7NO	6.4
Pyridine	Miscible	115.2	C_5H_5N	5.3
Methanol	Miscible	65	CH_3OH	5.1
Acetone	Miscible	56.3	$(CH_3)_2CO$	5.1
Chloroform	Slightly	61.2	$CHCl_3$	4.1
Dichloromethane	Slightly	39	CH_2Cl_2	3.1
1,2-dichlorobenzene	Poorly	181	$C_6H_4Cl_2$	2.7

* polarity increases with increasing polarity index

3.0　PROCEDURE

3.1　Paint Flakes Preparation

Approximately 5 mLs of household liquid paint was poured onto a metal plate (11 cm x 22 cm) and spread out into a thin layer using a paint brush. The painted

surface was left overnight to dry. The dried paint was gently scraped from the metal surface to obtain paint flakes. Paint flakes were made into a powder form using a mortar and pestle for subsequent IR tests.

3.2 Solubility Testing

100 mg of paint flakes were dissolved in 5 mL of each organic solvent in a test tube. The test tube was placed in a heated water bath for 5 mins The temperature of the water bath was set at $35°C$ for dichloromethane solvent and $50°C$ for the remaining solvents.

3.3 IR Spectroscopy

Paint flakes were directly subjected to ATR-FTIR spectroscopy under reflectance mode with no sample preparation. A small amount of paint flake was placed on the diamond cell and scanned for 32 times for good spectral collection.

3.4 System Operational Flow

The operational flow of the database system is designed according to the forensic paint analysis standard [5], whereby the physical appearance of paint sample is always the first parameter to be compared. The system will compare the physical characteristic of paint sample, followed by the chemical testing result, and finally the comparison of IR spectra.

For each step, the user will be given a clear instruction in either a pop-up message or static text box mode. There are 3 main steps in paint identification. Each of the steps has been divided into 3 different system pages. As the user move along the steps, possibility of the unknown to match with the known samples in the database will increase. This is due to the filtering features of the system. Each comparison step will filter out the unmatched sample and leave the matched one. From the filtered list of sample, comparison will proceed until the final stage, which is the IR spectrum matching page.

The household paint database system utilized Microsoft SQL Server 2008 as the Database Management System (DBMS), while the interface was developed using Microsoft Visual Studio 2008. All the data is stored in relational tables, mainly to increase data integrity and avoid data redundancies. The summary of the household paint database system flow is shown in Figure 1.

Figure 2 shows the color matching page, where different physical color of the paint available in the database is displayed for selection. The possible list of sample(s) that matched the color selected by user will be displayed. With the list of samples obtained, a user can proceed to the solubility test comparison, as shown in Figure 3. The user needs to click or "check" on the solvent box wherein sample is soluble. From there, the system will filter out the previous list. Only sample that matched the solubility test result will be kept for further matching in the next step. The unmatched sample will be excluded. The final list will move on to the **IR** page, where the remaining sample that is most relevant to the unknown will be kept for **IR** functional group comparison, as shown in Figure 4.
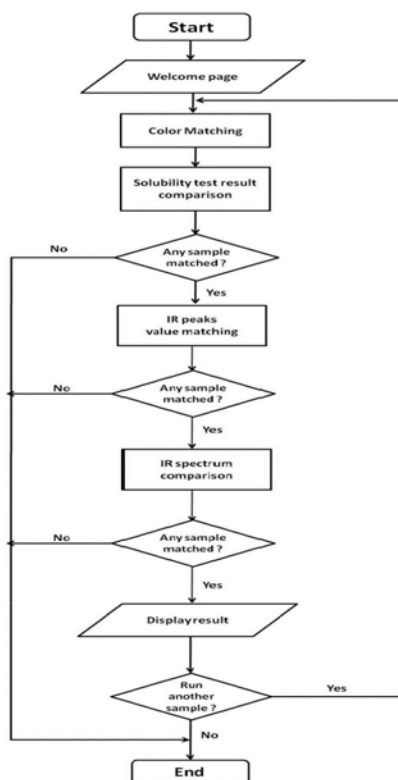

Figure 1 System flow for household paint database

Figure 2   Color matching page of the database system

Figure 3   Solubility result matching page of the paint database system

Figure 4 Final page showing IR result matching

4.0 RESULTS & DISCUSSION

4.1 Paint Information

A myriad of building paints are commercially available: Also available are specific paint types such as anti-bacterial, anti-fungal, spot-less, easy wash, odor-free and many more. Each type of paint coat offers a wide range of colors depending on the consumer's choice. In the physical examination of the paint sample, all important paint features were recorded for reference.

A total of eight parameters were selected, serving as data fields for building up the database system. The brand name, model number and paint category were obtained directly from the label of the paint can. The base type (water or solvent base) and paint usage information were obtained from the product specification sheet. Physical color appearance was determined from the paint flakes prepared, and not originally from the paint liquid. Remarks on field collected gave the additional information for the physical characteristic of the prepared paint flakes, for instance, the glossiness and transparency of the paint flake.

4.1.1 Sample Coding

Sample coding for a paint sample is important when dealing with experiments related to database building. A proper paint coding is effective for identifying the paint sample during the experimental phase. Apart from that, paint coding also acts as a primary key in the database building stage which is used for information storing. With this primary key, each sample becomes unique and easily identified by the system [6]. The coding consisted of 5 attributes, each separated by a hyphen, attributors are derived as:

Color & Appearance – Base Type – Usage – Brand – Model No

For a white glossy water-based paint sample, abbreviation W was used to denote the white color and GL for glossy and WB to denote water-based paint. **WGL-WB-I-S-8500** is an example of sample code for an actual white glossy paint sample. The details of coding used for the samples are shown in Table 2.

4.2 Data Types

The data types used for the database building consisted of quantitative, alphanumerical and graphical data. IR wave number values were considered as quantitative or numerical data and an IR spectrum categorized as graphical data. Paint information and chemical test results comprised of the alphanumerical data type [7].

4.2.1 Chemical Testing

A total of seven organic solvents were employed in the solubility test for all the paint samples analyzed. The color of the solution produced was recorded. Figure 5 displays the database storage view for solubility test result for several paint samples.

Table 2 Codings used for paint sample

Attributes	Name	Code
1. Observed color	Clear	CL
	White	W
	Yellowish White	YW
	Greenish White	GW
	Silver	S
	Metallic Silver	MS
	Grey	GR
	Light Grey	LGR
	Black	BK
	Red	R
	Dark Red	DR
	Light Red	LR
	Light Pink	LP
	Light Orange	LO
	Yellow	Y
	Greenish Yellow	GY
	Dark Yellow	DY
	Green	GN
	Light Green	LGN
	Blue	BL
	Dark Blue	DB
	Greenish Blue	GB
	Light Blue	LB
	Light Violet	LV
2. Appearance	Glossy	GL
	Transparent	T
	Semi-Transparent	ST
3. Base Type	Water-based	WB
	Oil-based	OB
	Solvent-based	SB
4. Usage	Interior	I
	Exterior	E
5. Brand name	Seamaster	S
	Nippon	N
	ICI Dulux	D
	Colorland	C

Tiger	T
Anchor	A
Advance	AD
Samurai	SM

Figure 5 Solubility test result incorporated in the database field records

4.2.2 IR Spectrum

The types of functional groups present in paint may differ and vary from one sample to another. The chemical compounds normally reflect on the functions and features that are provided by the paint.

Peak wave number values are important in the comparison and identification of unknown paint sample. It becomes the main criteria during the unknown case sample comparison [8]. However, only IR diagnostic peaks values were stored in the database, where a total of ten diagnostic peaks were selected. The selection of diagnostic peaks was based on the commonly appeared functional groups in

building paint samples. The absence or presence of the functional group provides a way for differentiating between samples which have similar physical appearance and solubility result. IR spectra of paint samples were also stored for manual visual matching.

For the information storing, absence of a diagnostic peak was assigned zero while peak values in wave number (cm^{-1}) were stored for specific functional groups that were present in a sample. Absence of a group commonly known as "no-band" observation is another important criterion during IR analysis. The determination of the presence a functional group requires a certain degree of knowledge and experience on analyzing the spectrum. This is due to the overlapping range for some functional group and variability of peak intensities. Table 3 summarizes the common functional groups found in building paint while Figure 6 shows wave number of the ten diagnostic peak values, sample code, and binary data of IR spectra.

Table 3 Common functional groups typically found in building paints [9]

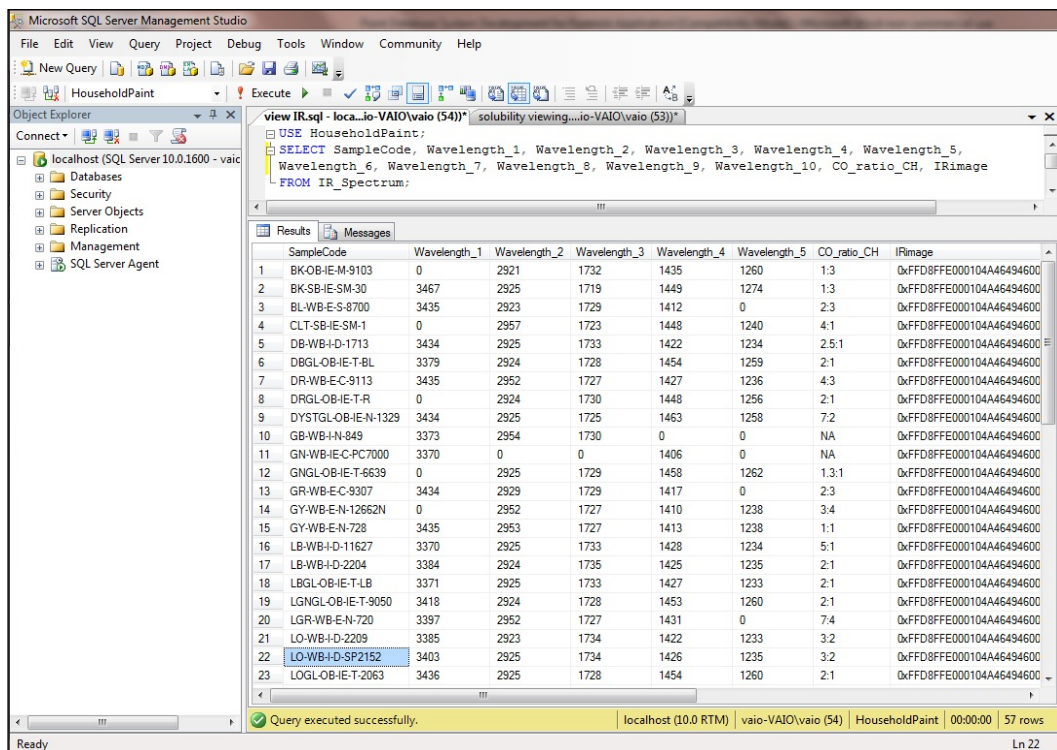| Functional group	Wavenumber (cm^{-1})
O-H stretch	3200-3600
C-H stretch	2850-3000
C=O stretch	1715-1740
C-H bend	1350-1480
C-O stretch	1210-1320
C-O stretch	1050-1170
C-O stretch	1000-1250
=C-H out-of-plane aromatic bend	870 - 880
=C-H out-of-plane bend	740 - 755
=C-H out-of-plane bend	695 - 715

Figure 6 Wavelengths of diagnostic IR peaks stored in the database

4.3 Data Retrieving

The main purpose of paint database system is to identify unknown paint sample by comparison with those stored in the paint database. The basic flow of the system starts from color matching, followed by solubility result comparison, and finally **IR** functional group matching. Since the results from ATR-FTIR spectroscopy and the **IR** analysis were not involved in any part of the comparison flow, the use of search algorithms can be avoided in order to reduce the complexity of the information searching [10].

The functional group matching and data retrieval were based on the absence and presence of ten diagnostic groups stored in the database. On the **IR** diagnostic group comparison page, the user will be required to select the presence of diagnostic group of the unknown case sample. The system gathers the inputs from

the user. Comparison is made for both user's input and the most related samples retrieved from the database.

The developed building paint database utilized a list filtering method. Since spectrum matching is the last stage for sample comparison, the initial prior stage, solubility result comparison, has filtered out the least related one. Possible sample size has greatly been reduced and the accuracy of matching the most related sample has consequently been increased. The system with minimized user action has highly increased the efficiency of data input and reduce human errors [11]. The overview of the data retrieving process is as summarized in Figure 7.
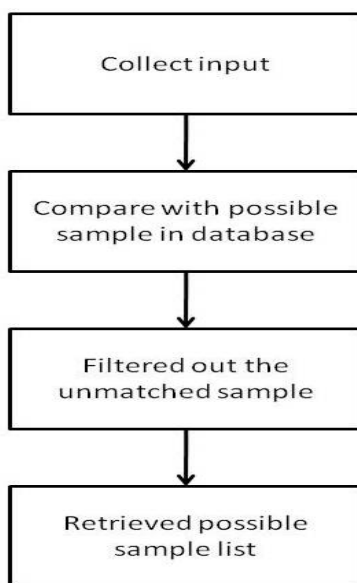

Figure 7   Flow chart for data retrieval

4.4 System Features

The main interface consists of two options: Sample matching page and administrator page. A user guide page is provided for the first-time user. A brief introduction on how the system works and the flow for unknown sample matching is included. On logging in, the user has to enter his/her name, identification card number, number of samples that are involved in sample matching, and the nature of the unknown samples. A record for each user will be generated as reference.

After collecting the user log details, the user will be directed to the color matching page. The complete system flow is as shown in Figure 1.

The administrator page consists of four different functions: view, insert, edit, and delete. The paint information, solubility test result and IR spectrum wave number include in each different functions. In other words, the administrator has the rights to view, insert, update or even delete the information for paint details, chemical testing result and also IR spectrum for each sample.

5.0 CONCLUSIONS

A database system for the forensic comparison of building paints has successfully been developed. Time taken for the whole matching process was relatively short as compared to manual matching. The developed building paint database utilized a list filtering method. The database system accuracy was considered high, since the system was capable of excluding the irrelevant samples by filtering out according to the chemical results inserted by the user. System features greatly influence the performance of the system. The developed paint database system offers an alternative and attractive means of comparison of unknown building paint samples particularly in forensic criminal cases where control paint samples are not found.

REFERENCES

[1]   Corte-Real, F. 2004. Forensic DNA Database. *Forensic Sci. Int.* 146S: S143-S144.

[2]   Automotive Paint Database, P. D. Q. Royal Canadian Mounted Police. 2007.

[3]   Thompson, W. C., F. Taroni, and C. G. G. Aitken. 2003. How the Probability of a False Positives Affects the Value of DNA Evidence. *J. Forensic Sci.* 48(1): 47-54.

[4]   Rumble, J. R. and D. R. Lide. 1985. Chemical and Spectral Database: A Look Into the Future. *J. Chem. Inf. Computer Sci*. 25: 231-235.

[5]   American Society for Testing and Materials (ASTM) International. 2008. *Standard Guide for Forensic Paint Analysis and Comparison*. USA, E1610-02.

[6]   Fahl, G. and T. Risch. 1997. Query Processing Over Object Views of Relational Data. *VLDB J*. 6: 261-281.

[7]   Masui, H. 1996. SPECTRA: A Spectral Information Management System Featuring a Novel Combined Search Function. *J. Chem. Inf. Computer. Sci*. 36: 294-298.

[8]   Warr, W. A. 1991. Spectral Databases. *Elsevier Sci. Pub*. 10: 279-292.

[9]   Bell, S. 2006. *Forensic Chemistry*. 1ª ed. Upper Saddle River: Pearson Education Inc.

[10]   Lowry, S. R., D. A. Huppler, and C. R. Anderson. 1985. Data Base Development and Search Algorithms for Automated Infrared Spectral Identification. *J. Chem. Inf. Computer Sci*. 25: 235-241.

[11]   Hueni, A., J. Nieke, J. Schopfer, M. Kneubuhler and K. I. Itten. 2009. The Spectral Database SPECCHIO for Improved Long-Term Usability and Data Sharing. *Computer & Geosciences*. 35: 557-565.