

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA – CT
DEPARTAMENTO DE COMPUTAÇÃO APLICADA
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**MÚLTIPLOS SISTEMAS DE RACIOCÍNIO BASEADO
EM CASOS PARA AUTENTICAÇÃO DE PINTURAS**

TRABALHO FINAL DE GRADUAÇÃO

Andreo de Aguiar Vieira

**Santa Maria, RS, Brasil
2013**

MÚLTIPLOS SISTEMAS DE RACIOCÍNIO BASEADO EM CASOS PARA AUTENTICAÇÃO DE PINTURAS

Andreo de Aguiar Vieira

Monografia apresentada ao Curso de Ciência da
Computação do Departamento de Computação Aplicada da Universidade
Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do
grau de
Bacharel em Ciência da Computação.

Orientador: Luís Alvaro de Lima Silva

**Trabalho de Graduação N° 359
Santa Maria, RS, Brasil
2013**

**Universidade Federal de Santa Maria
Centro de Tecnologia
Curso de Ciência da Computação**

A Comissão Examinadora, abaixo assinada,
aprova o Trabalho de Graduação

**MÚLTIPLOS SISTEMAS DE RACIOCÍNIO BASEADO EM CASOS
PARA AUTENTICAÇÃO DE PINTURAS**

elaborado por
Andreo de Aguiar Vieira

como requisito parcial para obtenção do grau de
Bacharel em Ciência da Computação

COMISSÃO EXAMINADORA



Luis Alvaro de Lima Silva, PhD
(Presidente/Orientador)



Ana Trindade Winck, Dr. (UFSM)



Lisandra Manzoni Fontoura, Dr. (UFSM)

Santa Maria, 25 de Julho de 2013

AGRADECIMENTOS

Dedico os meus agradecimentos a minha família, que me ajudou, apoiou e aconselhou durante esta fase da minha vida. Meu pai João Batista e minha mãe Adiles Vieira. Agradeço em especial a minha noiva, Karolina Simon, que a quatro anos torna a minha vida mais alegre, atura e me entende, especialmente nos momentos de stress. Gostaria de agradecer também a Mila e a Luly que são capazes de me alegrar em qualquer momento. Gostaria de agradecer ao professor Luís Alvaro pela sua orientação, permitindo que este trabalho fosse realizado. Agradeço também a todos os amigos que, de alguma forma, ajudaram durante este período na Universidade Federal de Santa Maria.

I would like to thank Dr. Nicholas Eastaugh for his kindness, providing the material that made this work possible.

RESUMO

Trabalho de Graduação
Curso de Ciência da Computação
Universidade Federal de Santa Maria

MÚLTIPLOS SISTEMAS DE RACIOCÍNIO BASEADO EM CASOS PARA AUTENTICAÇÃO DE PINTURAS

AUTOR: ANDREO DE AGUIAR VIEIRA

ORIENTADOR: PROF. DR. LUÍS ALVARO DE LIMA SILVA

Data e Local da Defesa: Santa Maria, 25 de julho de 2013.

O desenvolvimento de sistemas de Raciocínio Baseado em Casos (Case-Based Reasoning - CBR) é apropriado quando a solução para um problema pode ser encontrada a partir da recuperação e reuso de experiências de solução de problemas passados. Neste contexto, o objetivo deste Trabalho de Graduação (TG) é desenvolver técnicas de CBR para auxiliar o processo de autenticação de data de pinturas históricas. Em particular, o TG explora a aquisição automática ou semiautomática de dados e conhecimento provenientes de experiências de solução de problemas de autenticação, as quais são representadas como “casos” em um sistema web de CBR denominado AuthenticPainting. Além disso, o TG descreve como desenvolver a análise de similaridades entre diferentes tipos de atributos típicos de relatórios de autenticação de pinturas, os quais, neste TG, são: pigmentos históricos e composições elementares de pinturas. Em geral, este trabalho busca combinar e testar os resultados obtidos quando múltiplos mecanismos de recuperação de casos, ou múltiplos sistemas de CBR, baseados nestes atributos são utilizados na solução de novos problemas. Os resultados obtidos permitem concluir que a utilização de múltiplos sistemas de CBR, juntamente com o método de combinação testado para estes sistemas, é relevante para auxiliar as tarefas de autenticação de data de pinturas quando um dos sistemas de CBR envolvidos é capaz de gerar resultados que possam ser confirmados pela solução apresentada pelo outro sistema de CBR.

Palavras-chave: Inteligência Artificial. Raciocínio Baseado em Casos. Autenticação de Pinturas.

ABSTRACT

Trabalho de Graduação
Undergraduate Program in Computer Science
Universidade Federal de Santa Maria

COMBINATION OF MULTIPLE CASE-BASED REASONING FOR PAINTING AUTHENTICATION

AUTHOR: ANDREO DE AGUIAR VIEIRA

ADVISOR: PROF. DR. LUIS ALVARO DE LIMA SILVA

The development of Case-Based Reasoning (CBR) systems is suitable when the solution for a problem can be found from the retrieval and reuse of past experiences of problem solving. In this context, the main goal of this Undergraduation Dissertation is to develop CBR techniques to support the process of date authentication of historical paintings. In particular, this dissertation exploits the automatic or semi-automatic acquisition of data and knowledge coming from experiences of problem solving, which are represented as "cases" in a web-based CBR system called AuthenticPainting. Furthermore, this work describes how to develop the analysis of similarities between different types of typical attributes found in painting authentication reports, which in this dissertation are historical pigments and elemental compositions of paintings. In general, this dissertation aims to combine and test the results that can be obtained when multiple mechanisms of case retrieve, or multiple CBR systems, based on these attributes are used on the solution of new problems. The results obtained allow one to conclude that the use of multiple CBR systems, along with the combination method which was tested with these systems, is relevant to assist the authentication of painting dates when one of the involved CBR systems is able to produce results that may be confirmed by the solution presented by the other CBR system.

Keywords: Artificial Intelligence. Case-based Reasoning. Painting Authentication.

LISTA DE ILUSTRAÇÕES

Figura 2.1 – Ciclo do CBR	16
Figura 3.1.1 – Introdução de um caso de autenticação	23
Figura 3.1.2 – Exemplos de análises de uma pintura	24
Figura 3.1.3 – Exemplos de discussões de autenticação	25
Figura 3.1.4 – Conclusão de um caso de autenticação	25
Figura 4.1 – Funcionamento do AuthenticPainting	28
Figura 4.2 – Diagrama de Classes do sistema Authentic Painting	29
Figura 4.1.1 (A) – Inserção de elementos com arquivo CSV	31
Figura 4.1.1 (B) – Caixa de seleção do arquivo CSV	32
Figura 4.1.1 (C) – Input após o arquivo selecionado	32
Figura 4.1.2 – Modo de inserção manual de elementos	34
Figura 4.2.1 – Introdução PaintingC	35
Figura 4.2.2 – Conclusões PaintingA	36
Figura 4.2.3 – Conclusões PaintingB	37
Figura 4.2.4 – Resultado da recuperação utilizando composições elementares	39
Figura 4.3.1 – Base de pigmentos em arquivo XML	40
Figura 4.3.2 – Interface administrativa: pigmentos	40
Figura 4.3.3 – Inserção de pigmentos na aquisição de um caso	41
Figura 4.4.1 – Recuperação de casos similares utilizando pigmentos	43
Figura 4.6.1 – Confirmação da pintura PaintingB como a mais relevante para o processo de autenticação da pintura alvo PaintingC	46
Figura 5.1 – Percentagem de casos em diferentes períodos de tempo presentes nas bases de casos	47
Figura 5.1.1 – Resultados da consulta da pintura p75 utilizando pigmentos	50
Figura 5.1.2 – Resultados da consulta da pintura p75 composições elementares	50
Figura 5.1.3 – Resultados combinação utilizando peso 2 para os pigmentos	51
Figura 5.1.4 – Resultados das análises das recuperações por pigmentos	52
Figura 5.1.5 – Resultados das análises das recuperações por composições elementares..	53
Figura 5.1.6 – Resultados das análises com peso 3 para pigmentos	53

LISTA DE TABELAS

Quadro 4.1.1 – Exemplos de resultados EDX	32
Quadro 4.1.2 - Resultados EDX após o primeiro passo de codificação	33
Quadro 4.1.3 - Dados finais de elementos químicos do caso hipotético	33
Quadro 4.2.1 - Codificação dos resultados EDX para PaintingC	36
Quadro 4.2.2 - Resultados codificados para a pintura PaintingA	36
Quadro 4.2.3 - Resultados EDX codificados para a pintura PaintingB	37
Quadro 4.2.4 - Baseado em composições elementares, casos similares a PaintingC	38
Quadro 5.1.1 - Análise dos resultados da pintura p75 para pigmentos	49
Quadro 5.1.2 - Análise dos resultados da pintura p75 para composições elementares ...	50
Quadro 5.1.3 - Análise dos resultados da pintura p75 utilizando peso 3 para os pigmentos	51

LISTA DE ANEXOS E APÊNDICES

Anexo A – Resultados do teste do sistema	58
Apêndice A – Um relatório de autenticação de pinturas	78

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Objetivos	12
1.2 Uma área de pesquisa e aplicação para este Trabalho de Graduação: Autenticação de Data de Pinturas Históricas	12
2 RACICÍNIO BASEADO EM CASOS	15
2.1 Aquisição e representação de casos	16
2.2 Recuperação de casos (análise de similaridade)	17
2.2.1 Algoritmo da vizinhança	17
2.3 Retenção de casos	18
2.4 Teste em CBR	18
2.5 Sistemas de CBR com múltiplas bases de casos	19
3 AUTENTICAÇÃO DE PINTURAS	21
3.1 A Representação de um caso de autenticação pinturas	22
4 ANÁLISE, PROJETO E IMPLMETAÇÃO DO CBR AUTHENTICPAINTING	27
4.1 Exemplo de um caso de autenticação de pinturas: representação e codificação de um caso utilizando composições elementares	30
4.2 Processo de recuperação de casos utilizando composições elementares	34
4.3 Exemplo de caso: representação de um caso utilizando pigmentos	39
4.4 Processo de recuperação de casos utilizando pigmentos	41
4.5 Problema de Combinação de Datas em Autenticação de Pinturas	43
4.6 Processo de combinação dos resultados gerados por pigmentos com os resultados gerados por composições elementares	45
5 TESTES DO SISTEMA	47
5.1 Um exemplo de teste do sistema utilizando a pintura p75From1880ToEarly20th como consulta	49
6 CONCLUSÕES E TRABALHOS FUTUROS	54

1 INTRODUÇÃO

Encontrar a solução para um determinado problema pode ser uma tarefa complexa. Em muitas situações, contudo, esta tarefa pode ser realizada por uma pessoa com um vasto conhecimento e experiência na solução de problemas em um determinado domínio de aplicação. A esta pessoa damos o nome de especialista.

Um especialista em um assunto particular acumula durante a sua vida profissional diversas situações relevantes de soluções de problemas. Este mesmo especialista, ao se deparar com um novo problema, usa a sua experiência passada buscando identificar problemas resolvidos anteriormente, os quais sejam similares ao novo problema a ser resolvido. Assim, a partir da experiência anterior, é possível chegar a uma maneira de solucionar o novo problema reutilizando e adaptando uma das soluções alcançadas no passado. Este comportamento de solução de problemas centrado no reuso de experiências passadas pode ser exemplificado pela maneira que um médico realiza um diagnóstico. Por exemplo, o médico pode, através da identificação dos sintomas de um paciente corrente, prescrever determinado medicamento para o paciente. Neste caso, os sintomas identificados no paciente corrente devem ser semelhantes aos do paciente anterior, sendo que este paciente anterior foi tratado com êxito quando o tratamento e medicamento em questão foi empregado.

Em computação, a área que focaliza o emprego de experiências passadas na solução de novos problemas é conhecida como Raciocínio Baseado em Casos (*Case-Based Reasoning - CBR*) (KOLODNER, 1993; LOPEZ DE MANTARAS et al., 2006). CBR é um paradigma de construção de sistemas inteligentes da Inteligência Artificial (IA). Este paradigma emprega técnicas computacionais para adquirir, explorar e reutilizar o conhecimento proveniente de experiências concretas de solução de problemas em diferentes domínios de aplicação (CHOU, 2009; KHOSHGOFTAAR; SELIYA; SUNDARESH, 2006; NAKASUMI, 2003). Em CBR, tais experiências são capturadas e representadas por um conjunto de “casos”.

Em um caso, diferentes conjuntos de características do problema podem ser importantes na determinação de uma solução. Neste cenário, múltiplos sistemas de CBR podem ser planejados e construídos assim como demonstrado em (LEAKE; SOORIAMURTHI, 2001). A ideia geral é que cada um destes sistemas pode buscar encontrar uma solução para o problema a partir da seleção e emprego de um conjunto de características particulares. Em princípio, a utilização de sistemas CBR fundamentados em diferentes características de um caso pode ajudar o especialista a explorar mais amplamente o espaço de possíveis soluções para um

problema dado, assim aumentando as possibilidades de obter uma solução melhor. Em geral, espera-se que tal melhoria de qualidade da solução apresentada possa ser alcançada através da combinação dos resultados gerados quando consultas são executadas em cada um dos sistemas de CBR disponíveis. Outra maneira de utilizarmos múltiplas bases de casos é quando uma destas gera resultados que são utilizados a fim de confirmar os resultados gerados por outra base de casos que, em determinado domínio de aplicação, sejam mais relevantes na obtenção das soluções. Na prática, entretanto, para que seja possível combinar ou confirmar estes resultados, cada sistema de CBR deve gerar soluções que obedecem a um padrão previamente estabelecido, o qual deve ser comum aos diferentes sistemas de CBR disponíveis.

1.1. Objetivos

1.1.1. Objetivos específicos

Este Trabalho de Graduação (TG) visa desenvolver e explorar técnicas voltadas para a aquisição sistemática de conhecimento na forma de casos. Além disso, dado um novo problema a ser resolvido e a partir de resultados obtidos através da recuperação de casos similares para uma consulta no sistema, o TG apresenta um método de combinação de resultados de recuperação provenientes da utilização de múltiplos sistemas de CBR.

1.1.2. Objetivos gerais

A ideia da técnica proposta é apresentar formas de combinar os resultados de recuperação de casos com o intuito de buscar obter uma solução mais compreensiva para um novo problema a ser resolvido pelo sistema. Por fim, o TG é fundamentado em uma área de pesquisa e desenvolvimento onde a solução de problemas é baseada casos concretos de solução de problemas: autenticação de data em pinturas históricas (EASTAUGH, 2006).

1.2 Uma área de pesquisa e aplicação para este Trabalho de Graduação: Autenticação de Data de Pinturas Históricas

O TG explora o problema de autenticação de data de pinturas históricas (EASTAUGH, 2006) como domínio de aplicação principal de pesquisa e desenvolvimento. A autenticação de uma pintura é uma tarefa desenvolvida por especialistas neste domínio de aplicação. No intuito de difundir este conhecimento especializado, bem como reduzir altos custos inerentes ao

processo de autenticação, a construção de técnicas que auxiliem a autenticação de pinturas históricas é, portanto, uma importante contribuição para esta área de aplicação e pesquisa por si só.

A pesquisa desenvolvida neste TG utiliza uma base de relatórios de autenticação de pinturas pré-existentes, ou casos, os quais foram desenvolvidos por um especialista neste domínio de aplicação. Esta base de casos é composta por 77 casos concretos de autenticação. Partindo desta base de casos, novos casos de autenticação podem ser adquiridos utilizando os métodos de aquisição de casos desenvolvidos no TG. Em geral, o TG propõe uma forma de padronização para o processo de construção destes relatórios e, conseqüentemente, sistemática ampliação da base de casos do sistema.

Neste TG, o processo de autenticação de data de pinturas baseado em CBR utiliza duas características ou propriedades básicas de pinturas: a) os pigmentos presentes na pintura e b) a composição elementar da pintura (os elementos químicos representando as misturas de pigmentos utilizadas pelo autor da pintura). Essas características são relevantes no processo de identificação do período histórico que determinada pintura pode pertencer (EASTAUGH, 2006; SILVA, 2010; SILVA et al., 2008, 2010). A determinação da data de criação de uma pintura é um dos fatores mais importantes no processo de autenticação de pinturas visto que esta investigação pode ser realizada a partir de métodos científicos variados, os quais podem ser combinados com o estudo das características artísticas que definem um determinado quadro e autor.

Assim como descrito por (SILVA, 2010), o emprego de técnicas de CBR podem gerar resultados utilizando cada uma dessas características (composições elementares e pigmentos) separadamente. Logo, este TG expande essa linha de pesquisa pela discussão de uma forma de combinar os resultados obtidos quando sistemas de CBR baseados nas duas características citadas são usados, permitindo que os resultados baseados em pigmentos possam ser confirmados pelos casos resultantes da recuperação baseada em composições elementares.

No Capítulo 2 são apresentados conceitos básicos de CBR, o ciclo utilizado pelos sistemas de CBR, bem como técnicas para testes destes sistemas. Além disto, este capítulo revisa conceitos que envolvem a integração de múltiplos sistemas de CBR.

No Capítulo 3 é revisado o processo de autenticação de pinturas, identificado atributos importantes para a confirmação da data de uma determinada pintura. Além disto, o capítulo apresenta técnicas utilizadas pelo especialista em autenticação, a fim de obter os dados que são utilizados pelo sistema de CBR desenvolvido neste TG.

No Capítulo 4 é apresentado o sistema AuthenticPainting. Neste capítulo, também é apresentada a representação dos casos usada em cada um dos sistemas de CBR desenvolvidos, a qual é ilustrada por um exemplo básico de funcionamento para cada um destes sistemas. Por fim, este capítulo relata os problemas enfrentados devido a falta de padrões de representação de soluções e de pigmentos nesta aplicação. Por consequência, o capítulo apresenta padrões desenvolvidos que visam possibilitar a construção do sistema de CBR.

No Capítulo 5 é apresentada a metodologia de teste adotada neste TG. Neste capítulo, também é apresentado um exemplo completo de um teste, juntamente com uma análise crítica dos resultados obtidos. Por fim, os resultados dos testes realizados são apresentados, onde é possível avaliar a utilidade do sistema para o processo de autenticação de data de pinturas utilizando composições elementares e pigmentos.

Por fim, no Capítulo 6 são apresentadas as conclusões do TG e possíveis trabalhos futuros a serem desenvolvidos.

2 RACIOCÍNIO BASEADO EM CASOS

Raciocínio baseado em casos (*Case-Based Reasoning* - CBR) (KOLODNER, 1993; WATSON, 2003) é uma área da Inteligência Artificial (IA) que investiga como resolver problemas pela reutilização de soluções de problemas passados na solução de um novo problema. Em CBR, um problema presente na base de conhecimento do sistema é chamado de “caso”. Em essência, um caso é a contextualização da solução de um problema real que geralmente ensina alguma lição para a resolução de um novo problema (KOLODNER, 1993). Em CBR, a solução para um novo problema é principalmente buscada por meio da análise de similaridade entre casos: um caso corrente tomado como consulta no sistema e uma base de casos passados de solução de problemas.

O ciclo de CBR contendo recuperação, reutilização, revisão e retenção representa o funcionamento geral de um sistema de CBR (LOPEZ DE MANTARAS et al., 2006). A Figura 2.1 apresenta o modelo, adaptado de (AAMONDT; PLAZA, 1994), com suas quatro tarefas principais:

- Recuperação dos casos mais similares para um determinado caso usado como consulta no sistema;
- Reutilização da solução do caso mais similar (ou casos mais similares) na solução do novo problema;
- Revisão da solução, adaptando a solução para o contexto atual (visto que o novo caso pode conter diferenças em relação ao caso recuperado pelo sistema);
- Retenção do caso e de sua solução revisada para uma futura reutilização, assim buscando aprender como resolver novos problemas (porventura ampliando a competência de solução de problemas do sistema).

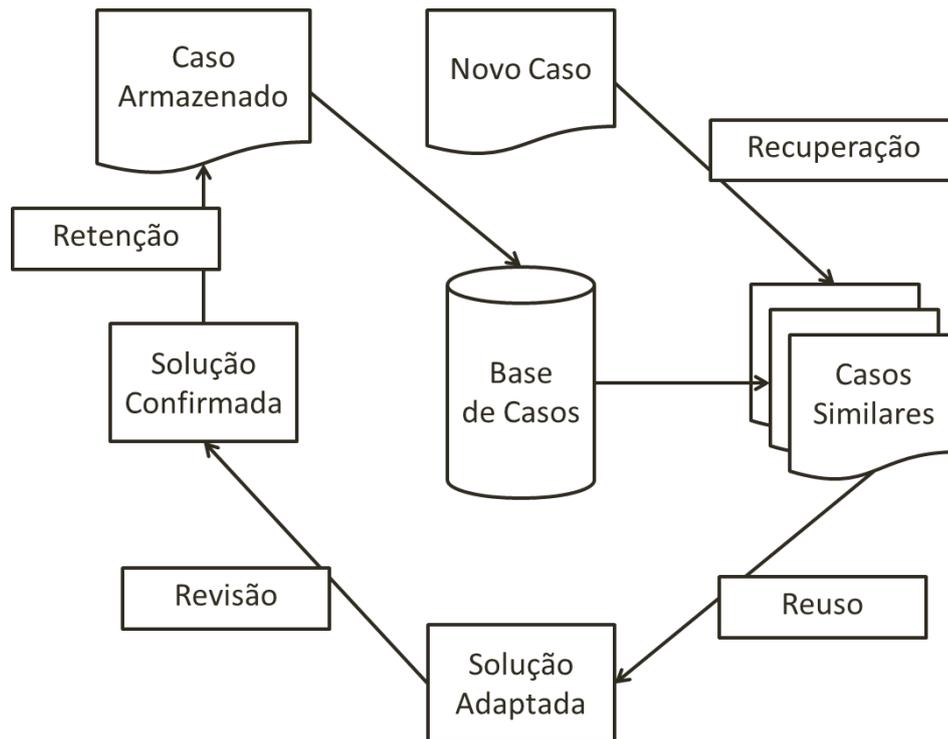


Figura 2.1 – Ciclo do CBR

2.1. Aquisição e representação de casos

Em muitas aplicações, o conhecimento utilizado em um sistema de CBR é proveniente de um especialista. Desta forma, para viabilizar a construção sistemática de sistemas de CBR é necessário desenvolver e empregar métodos automatizados ou semi-automatizados de aquisição de conhecimento, ou aquisição de casos. Em geral, métodos de aquisição de casos visam adquirir essas experiências anteriores e representar este conhecimento de uma maneira computável, assim permitindo estruturar o conteúdo destes casos para que seja possível avaliar a similaridade entre casos.

A tarefa de obtenção de casos possui diversos métodos simplificados de ser realizada, tais como reuniões e formulários, por exemplo. Obter os casos utilizando formulários permite que estes casos sejam estruturados conforme o padrão descrito pelo formulário. Além disso, esta técnica pode permitir seguir um vocabulário padrão de descrição que será utilizado na construção do sistema de CBR.

Uma vez que os casos sejam coletados e armazenados em uma base de casos, a representação destes segue um formalismo de representação de conhecimento: a) representação de um vetor de características, b) representação estruturada de casos (por exemplo, uma

representação orientada a objetos) e c) representação de casos em um formato semi-estruturado ou mesmo textual. A estrutura completa de um caso consiste, em geral, de cinco partes:

- Uma situação e seu objetivo;
- A solução e, às vezes, maneiras de alcançar ou derivar esta solução
- O resultado da utilização desta solução
- Explicação dos resultados alcançados nesta experiência de solução de problemas
- Lições que podem ser aprendidas a partir desta experiência.

2.2. Recuperação de casos (análise de similaridade)

Um sistema de CBR, ao se deparar com um novo problema ou consulta, busca em sua base de experiências os casos que sejam similares ao novo problema. Para que esta busca possa ser realizada, a base de casos precisa estar organizada e os casos devem estar indexados.

Encontrar casos similares a um novo problema é uma tarefa complexa, tendo em vista que diversos fatores podem ser significativos para que o sistema possa afirmar que dois casos são realmente similares. No entanto, a análise de similaridade é uma das tarefas mais importantes no desenvolvimento de sistemas de CBR, sendo necessário o desenvolvimento de algoritmos de recuperação eficientes nos diversos domínios de aplicação que as técnicas de CBR possam ser aplicáveis.

Um dos métodos mais importantes na recuperação de casos em CBR é o “algoritmo da vizinhança” (YU; JI; ZHANG, 2002). Neste método, as características de dois casos distintos são comparadas, gerando uma média ponderada dos resultados obtidos nessa comparação. Além disso, este método permite que diferentes características possuam relevâncias distintas na computação da similaridade entre os casos.

2.2.1. Algoritmo da Vizinhança

O algoritmo da vizinhança mais conhecido como KNN (do inglês K-nearest neighbors) (BISHOP, 2006; DE MANTARAS et al., 2001; YU; JI; ZHANG, 2002) busca identificar os k casos mais similares a um caso alvo (caso usado como consulta em sistemas CBR). Para o cálculo da similaridade entre dois casos, o primeiro passo é calcular a distância local entre cada uma das características tomadas individualmente. O cálculo da distância entre duas

características (e escalonamento/normalização dos resultados) é efetuado utilizando a seguinte equação:

$$dist(va_i, vb_i) = \left| \frac{va_i - vb_i}{max_i - min_i} \right| \quad (1)$$

Onde:

- va_i representa o valor da i -ésima característica do caso a;
- vb_i representa o valor da i -ésima característica do caso b;
- max_i representa o valor máximo da i -ésima característica;
- min_i representa o valor mínimo da i -ésima característica.

Uma vez que o cálculo de similaridades locais é realizado, é possível calcular a similaridade global entre dois casos. Para isso, diferentes métricas de distância podem ser usadas, onde a distância Euclidiana é a mais usada em muitas aplicações:

$$sim(a, b) = 1 - \sqrt{\frac{\sum_{i=1}^n (dist(va_i, vb_i))^2 * weight_i}{\sum_{i=1}^n weight_i}} \quad (2)$$

Onde:

- $weight_i$ representa o peso da característica i ;

2.3. Retenção de casos

Quando um sistema CBR encontrar a solução de um novo problema, este sistema pode armazenar este novo caso na base de casos. Normalmente, isso acontece quando esse novo caso possui características que o diferenciem de maneira significativa dos demais casos existentes na base de casos do sistema (ou seja, o novo caso descreve uma experiência que vale a pena ser guardada para futuro reuso). Em geral, o sistema deve adicionar apenas casos que possam ser úteis para futuras recuperações e consequente solução de novos problemas. Assim, a etapa de retenção de novos casos na base de casos permite ao sistema de CBR “aprender” como solucionar novos problemas, visto que a sua base de experiências é expandida.

2.4. Teste em CBR

Para garantir que o sistema de CBR seja eficiente (por exemplo, alcance os resultados esperados), é necessário realizar testes nos principais componentes do sistema. Dentre os métodos de teste disponíveis temos o método *Leave one out and test*. Tal como descrito em (WATSON, 2003), este método consiste em remover um dos casos presentes na base de casos do sistema e usá-lo como entrada/consulta no sistema.

Outro método importante de teste é o *Leave one in and test*, tal como descrito em (WATSON, 2003). De forma semelhante à técnica anterior, seleciona-se um caso existente da base de casos e, sem retirá-lo desta base, o utilizamos como entrada na consulta de casos similares. Com a utilização deste método é possível testar, principalmente, o algoritmo de recuperação de casos similares. Como resultado esperado deste teste, temos que o algoritmo de recuperação deverá encontrar o mesmo caso usado na entrada da consulta.

2.5. Sistemas de CBR com múltiplas bases de casos

A abordagem mais comum para a construção de sistemas de CBR consiste na utilização de um agente (neste caso, o sistema em si) que utiliza uma única base de casos (PLAZA; MCGINTY, 2005). Logo, isto nos faz acreditar que um sistema de CBR que originalmente possui os seus casos em diferentes bases de casos deve combinar esses casos em uma única base de casos. Neste caso, a utilização de múltiplas bases de casos pode ser aplicada em um sistema de CBR onde o sistema como um todo não dependa obrigatoriamente de todas as bases de casos envolvidas. Por exemplo, durante a recuperação dos casos similares, o sistema pode utilizar diferentes bases de casos para complementar a solução de um novo problema. Entretanto, o processo de combinação de múltiplas bases de casos pode acarretar em problemas, assim como descrito por (LEAKE; SOORIAMURTHI, 2002). Entre tais problemas, é possível citar:

- Disponibilidade dos casos: existe algum tipo de restrição ao acesso à base de casos (uma base de casos proprietária, por exemplo)
- Eficiência da combinação: combinar múltiplas bases de casos em uma única pode acarretar na perda de eficiência dos sistema, aumentando demasiadamente o tamanho da base de casos, onde podem estar sendo armazenados dados redundantes e/ou desnecessários
- Padronização da representação: por se tratarem de bases de casos diferentes, é muito provável que o padrão de representação dos casos também seja diferente (determinada

característica pode ser mais relevante em uma base de casos do que em outra, por exemplo)

- Manutenção da base de casos: com uma base de casos unificada, a manutenção da mesma pode ser problemática. Por exemplo, o que fazer quando a base de casos original é atualizada? Logo, isto poderá gerar, em alguns casos, uma base de casos que rapidamente perderá a fidelidade nos seus dados

A utilização de um sistema de CBR com múltiplas bases de casos pode permitir que as bases de casos sejam menores (focalizem determinado aspecto do problema e, por consequência, possuam um menor número de casos). Este fato pode vir a aumentar a eficiência do processo de recuperação e solução de novos problemas e, ao mesmo tempo, tornar os requisitos de armazenamento de casos mais gerenciáveis, assim como descrito em (LEAKE; SOORIAMURTHI, 2002).

Em resumo, um determinado problema de aplicação pode possuir múltiplas bases de casos disponíveis para serem utilizadas. Neste caso, sistemas de CBR que utilizem estas múltiplas bases de casos podem ser construídos seguindo as seguintes estratégias:

- Confirmar uma solução: em um determinado domínio, a utilização de múltiplas bases de casos pode ser importante quando uma das bases de casos envolvidas é capaz de gerar resultados mais relevantes que as demais bases de casos. Neste cenário, as bases de casos menos relevantes podem vir a confirmar ou reforçar os resultados de consultas obtidos quando a base de casos mais importante é utilizada no sistema;
- Criticar uma solução: de maneira semelhante a estratégia anterior, múltiplas bases de casos podem ser importantes quando o resultado gerado por uma destas bases é criticado ou refutado pelas demais bases de casos empregadas no sistema;
- Utilizar de maneira igualitária: neste caso, o sistema de CBR considera igualmente todas as bases de casos envolvidas (não existe uma base de casos que gere resultados mais relevantes para o problema a ser tratado).

3 AUTENTICAÇÃO DE PINTURAS

A autenticação de pinturas históricas consiste na análise da pintura com o objetivo de verificar se existem motivos para duvidar da: a) data atribuída à pintura e b) o artista que criou a pintura. Realizar a autenticação de autoria de uma pintura, apesar de possível, é uma tarefa complexa. Neste caso, a determinação de autoria está vinculada a diversos fatores históricos e proveniência que atualmente não podem ser comprovados por meios puramente científicos. Em contrapartida, identificar a qual período da história pertence uma pintura é um objetivo que pode ser alcançado em muitos problemas neste domínio de aplicação.

Segundo (EASTAUGH, 2006), a data ou período que pertence determinada pintura pode ser analisada através da observação de duas características principais: a) os pigmentos presentes na pintura (EASTAUGH, 2008) e b) a composição elementar da pintura (GOLDSTEIN et al., 2003) (os elementos químicos que compõem as misturas de pigmentos utilizadas por autores para produzir as diferentes cores usadas na construção de um quadro). Para que seja possível obter estas informações, amostras de diferentes áreas das pinturas são retiradas. Estas amostras devem ser oriundas de regiões da pintura as quais o especialista acredita que sejam da pintura original. Por exemplo, tais regiões não devem apresentar alterações significantes em relação às áreas adjacentes na pintura, tais como alterações oriundas de tarefas de restauração. No processo de autenticação de uma pintura, identificar áreas originais da pintura é extremamente importante, especialmente pelo fato de que devido ao desgaste provocado pelo tempo, muitas obras passam por processos de restauração. Além disto, as localizações escolhidas para retirada das amostras leva em consideração a amostragem do maior número possível de cores da pintura, sendo esta uma tentativa de identificar um maior número de pigmentos presentes no quadro.

a) Os pigmentos presentes na pintura

Conforme (EASTAUGH, 2006), os pigmentos encontrados em uma pintura possuem uma maior relevância no processo de investigação e atribuição da data. Neste caso, muitos pigmentos relevantes para o processo de autenticação de data têm um período conhecido em que foram utilizados na construção de quadros. Neste caso, tais pigmentos são inicialmente descobertos. Em seguida, eles começam a aparecer em pinturas desenvolvidas por diferentes autores e em diferentes locais. Depois, os mesmos passam a ser amplamente usados por diferentes autores. Finalmente, pigmentos começam a ser substituídos e eventualmente desaparecem por diferentes motivos. Logo, um especialista em autenticação de pinturas pode

afirmar que uma pintura não pertence a um determinado autor (ou seja, ela é falsa) se conseguir identificar um pigmento que não estava disponível para o autor na data originalmente atribuída à pintura. Em muitos casos, esta data é originalmente proposta e tomada como objetivo no processo de autenticação.

b) A composição elementar da pintura

Para a identificação da composição química de amostras de tinta e dos pigmentos presentes em uma pintura, diversos tipos de análises são realizadas em laboratório. O desenvolvimento ou não destas análises depende dos objetivos de autenticação almejados e dos custos envolvidos neste processo. A composição química de amostras de tinta de uma pintura pode ser obtida através da utilização de uma técnica conhecida como *Energy Disperse X-ray Spectrometry* (EDX) (GOLDSTEIN et al., 2003). É importante observar, contudo, que os pigmentos são determinados através dos resultados das análises químicas e de outras análises microscópicas realizadas, bem como pelo emprego de experiência prévia (baseada no conhecimento e perícia de especialistas na área) na identificação destes elementos de autenticação.

3.1. A Representação de um caso de autenticação pinturas

Um relatório de autenticação está dividido em “seções”, onde cada seção contém informações sobre a pintura e o processo de autenticação desenvolvido. Em geral, as principais seções de um caso de autenticação são a) introdução, b) análise, c) discussão e d) conclusões.

No problema de autenticação de pintura, um caso é normalmente formado por três partes que estão distribuídas nas diferentes seções de um relatório de autenticação. Em um caso de autenticação, o especialista apresenta as seguintes informações:

- A situação e o seu objetivo. Neste caso, a data atribuída que se deseja autenticar
- A solução encontrada. Neste caso, a descrição de uma ou mais datas para a pintura, as quais puderam ser obtidas no processo de autenticação
- A explicação dos resultados. Na discussão destes resultados, o especialista apresenta explicações a respeito dos resultados da análise. Tais explicações visam suportar a conclusão obtida na autenticação

A “introdução” de um relatório de autenticação visa identificar informações sobre a identificação da pintura (por exemplo: *PaintingD*), o autor que se acredita que tenha criado a

mesma (por exemplo: *Author of PaintingD*) e a data atribuída à pintura. Neste caso, o processo de autenticação vai tentar validar esta data inicialmente atribuída (por exemplo: *At 1910*). Fragmentos típicos da introdução de um caso real de autenticação podem ser vistos na figura 3.1.1 (A e B).

*(A) This report presents the results of a scientific examination of the painting **PaintingD**, thought to be by **Author of PaintingD** and to date from 1910. The painting measures 50cm (h) by 70cm (w) and has been executed on canvas.*

*(B) From the analyses and comparative studies, no pigments anachronistic for a date of 1910 were found. Moreover two pigments specifically identified as present in the painting are likely to be positively indicative of a work from the supposed date of the painting. It was also found that the types of pigments used are very similar to those that **Author of PaintingD** is known to have employed in the period.*

Figura 3.1.1 – Introdução de relatório de autenticação

A “análise” de um relatório de autenticação visa descrever o conjunto de técnicas utilizadas pelo especialista durante o processo de autenticação. Além disto, a análise apresenta informações visuais básicas da pintura (o tipo de tela da pintura, tipo de madeira utilizado na moldura, por exemplo). Contudo, informações importantes existentes nesta seção dizem respeito à composição elementar das amostras de tinta retiradas da pintura. Além disso, esta seção apresenta a identificação dos diferentes pigmentos presentes na pintura. Esta informação é obtida através do resultado das diferentes análises realizadas as quais são tomadas em conjunto com a experiência passada do especialista. Um exemplo típico da análise de uma pintura pode ser visto na Figura 3.1.2 (A e B).

Um dos problemas enfrentados na representação de um caso é a falta de padronização na descrição dos pigmentos identificados (ou seja, na nomenclatura usada para descrever os pigmentos identificados). Muitas vezes, esta falta de padronização está relacionada a inerente dificuldade de precisão que pode ser alcançada no processo de identificação dos pigmentos. Logo, a definição de um padrão de representação de pigmentos que abranja o maior número possível de pigmentos se faz necessário: este TG contribui para área de autenticação através desta padronização, e conseqüente construção de um sistema que permite os usuários empregar essa padronização de forma sistemática. Este padrão será apresentado na seção 4.3 deste TG.

(A) Prior to sampling, the painting was examined in detail visually, with the aid of low-power binocular microscopy, and with UV fluorescence. Little evidence of restoration was apparent, and the painting... The Painting appears to be still mounted on its original 4-piece adjustable wooden stretcher.

(B) The following pigments were identified: synthetic ultramarine, iron(III) hexacyanoferrate(II) ('Prussian blue'), manganese phosphate ('manganese violet'), 'dry' process Mercury(II) sulfide ('vermillion'), copper di-potassium hexacyanoferrate(II) ('Hatchett's brown' and other terms)...

Figura 3.1.2 – Exemplos de análise de uma pintura

A “discussão” de um relatório de autenticação apresenta diferentes observações, explicações e justificativas (ou argumentos) à respeito dos resultados obtidos durante o processo de análise da pintura. Em muitas discussões, estes argumentos focalizam a identificação e caracterização de pigmentos chave para o processo de autenticação. Quando encontrados, estes pigmentos possuem uma maior importância na autenticação da data de uma pintura. Em resumo, um dos principais objetivos da “discussão” é analisar os pigmentos encontrados, buscando verificar se estes pigmentos estavam disponíveis para um determinado autor da pintura na data atribuída a pintura. Para isto, o especialista utiliza a sua experiência passada na autenticação de pinturas juntamente com o conhecimento existente sobre pigmentos utilizados historicamente pelos mais diversos artistas (EASTAUGH, 2008). Além disto, o especialista discute se os pigmentos encontrados na pintura estão consistentes com os que normalmente foram utilizados pelo artista em questão (normalmente, por meio de uma análise comparativa entre pinturas desenvolvidas por um determinado artista).

Exemplos de discussões são apresentados na figura 3.1.3 (A e B) através de fragmentos de um relatório de autenticação.

(A) No pigments inappropriate for a date of 1910 were found. For example, numerous false paintings in the style of the Russian Avant Garde movement contain titanium dioxide white pigments, for which current evidence suggests there was little general use before 1950s. This pigment was not detected, not other like it.

(B) Detection of manganese phosphate is important to the probable dating of the painting. Several manganese phosphates are listed in the historical pigment literature, notably ammonium manganese phosphate $((\text{NH}_4)_2\text{Mn}_2(\text{P}_2\text{O}_7)_2)$, and manganese(III) phosphate. ... with a number of sources indication the ammonium form was that typically produced; it seems probable though that both form were commercially available at that time.

Figura 3.1.3 – Exemplos de discussões de autenticação

Por fim, as “conclusões” de um relatório de autenticação expõem as principais soluções sobre a data da pintura sendo autenticada. Tais soluções buscam apresentar e discutir as principais evidências para determinação de data alcançadas pelo especialista em autenticação de acordo com as análises realizadas. Neste caso, grande parte do processo de autenticação visa coletar e analisar evidências para determinar a data ou período de datas que pintura pode pertencer. Dentre outras situações, caso um pigmento identificado na pintura não estiver disponível na data proposta para a pintura, o especialista pode afirmar que a pintura trata-se de uma falsificação. Na figura 3.1.4, é apresentado um exemplo típico de conclusão de autenticação.

This analysis has shown that there are no overt reasons on the basis of the pigments in the painting to suppose that this Picture is other than it is currently supposed to be. Apart from the absence of anachronistic materials, the presence of a number of pigments consistent with early twentieth century forms and use supports this contention.

Figura 3.1.4 – Conclusão de um caso de autenticação

É importante observar que uma solução apresentada neste processo de autenticação de pinturas visa atender os objetivos de autenticação determinados pelos clientes (investidores em obras de artes, museus, etc.). Na prática, tais clientes requerem a execução destas tarefas de autenticação. Entre outros objetivos, a maioria dos processos de autenticação visam determinar se existe razão para uma pintura não ter sido desenvolvida por certo artista em certa data.

A data de uma pintura obtida a partir do processo de autenticação é geralmente expressa de diferentes formas por um especialista. Esta data também pode ser descrita com diferentes graus de detalhamento e precisão. Por exemplo, uma data pode ser descrita como “*Maybe 1914 Likely 1930 to 1960*”. Observe que cada data particular que uma solução venha a possuir apresenta uma importância distinta (por exemplo: *Likely 1930 to 1960*). Neste caso, dentre as conclusões de datas ou períodos de datas pertencentes a uma solução, um destes períodos pode ser mais provável que outro. Isso acontece devido ao fato do especialista ter encontrado evidências mais fortes que suportem um determinado intervalo de datas. Normalmente, esta probabilidade é descrita informalmente em relatórios de autenticação (ao invés de conter especificações numéricas de probabilidade). Em geral, a determinação de probabilidade de datas é baseada em uma combinação de todas as evidências encontradas e analisadas na pintura.

Na descrição de um caso de autenticação, esta especificação de confiança pode ser representada através de um campo *confidence* na classe *Dates* (ver figura 4.2 do Capítulo 4).

4. ANÁLISE, PROJETO E IMPLEMENTAÇÃO DO CBR AUTHENTICPAINTING

O AuthenticPainting é um sistema web resultante do estudos de relatórios textuais de autenticação de pinturas. Este sistema foi desenvolvido em Java utilizando o *framework* de desenvolvimento PlayFramework, o qual é específico para desenvolvimento web. A escolha de uma plataforma web vem de encontro a um dos principais objetivos deste projeto: o desenvolvimento de um sistema que possa padronizar os dados produzidos no desenvolvimento destas tarefas de autenticação. Além disso, o sistema busca apresentar uma forma de viabilizar a aquisição sistemática de novas experiências de autenticação, as quais são gravadas no formato de casos para CBR. A versão do PlayFramework escolhida é a versão 2.1.1 (a versão mais estável disponível até a conclusão deste TG). O Sistema de Gerenciamento de Banco de Dados (SGBD) PostgreSQL 9.2 foi utilizado para o armazenamento da base de casos. Neste banco de dados, outras informações utilizadas pelo sistema também são armazenadas, tais como a identificação dos usuários que podem utilizar o sistema, por exemplo.

Um caso de autenticação de pinturas é representado no sistema AuthenticPainting por um conjunto de classes em um modelo orientado a objetos. Neste caso, um especialista em autenticação de pinturas analisa as características de uma pintura e seus valores, observando propriedades químicas e pigmentos identificados que podem ser relevantes para a determinação da data de uma pintura. Na modelagem do sistema, essas informações representam diferentes classes, onde cada uma dessas classes possui características importantes. Desta forma, no sistema AuthenticPainting, um caso possui como principais características a) um conjunto de pigmentos e b) um conjunto de elementos químicos.

Assim como investigado em (SILVA, 2010; SILVA et al., 2008, 2010), pigmentos e composições elementares de uma pintura são as informações de entrada para o sistema CBR. Neste caso, estes dois conjuntos de propriedades distintas de pinturas são utilizados para a determinação de similaridades entre casos no sistema.

Nesta aplicação de autenticação de pinturas, os dados sobre composições elementares estão geralmente disponíveis no formato de planilhas eletrônicas. O restante das informações de um caso pode ser obtido pela análise de relatórios textuais de autenticação gerados pelo especialista na área. Os dados de identificação da pintura e o autor que possivelmente criou a pintura, além de informações sobre o intervalo de datas (ou data) atribuídas a pintura, estão

disponíveis logo na introdução destes relatórios. Um exemplo de funcionamento do sistema AuthenticPainting é apresentado na figura 4.1.

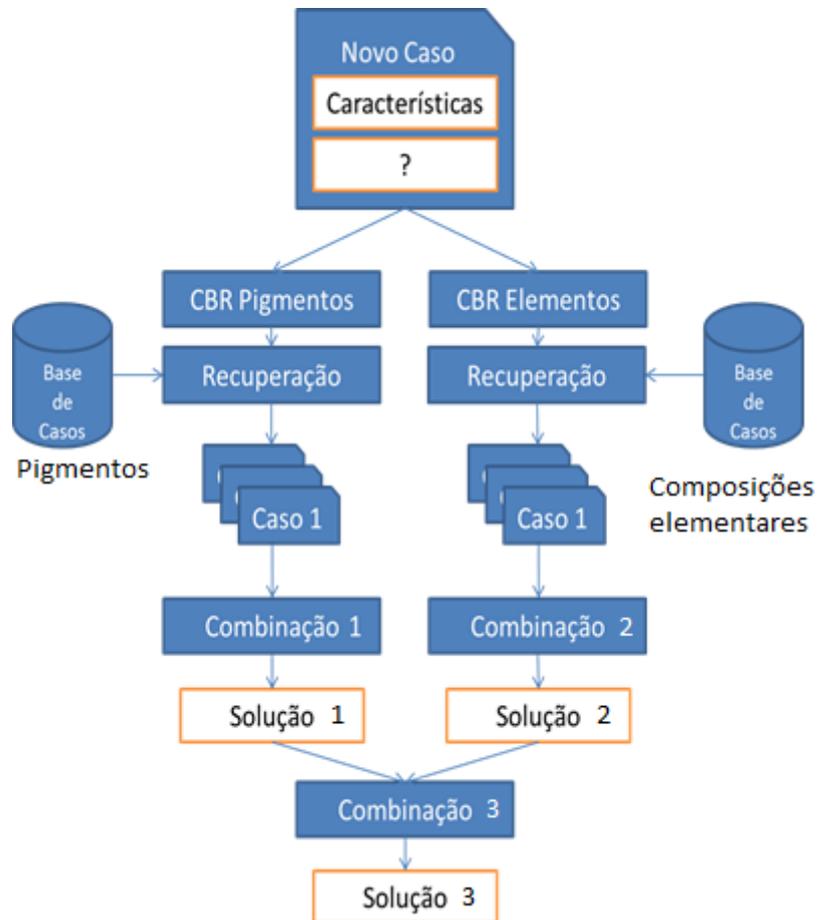


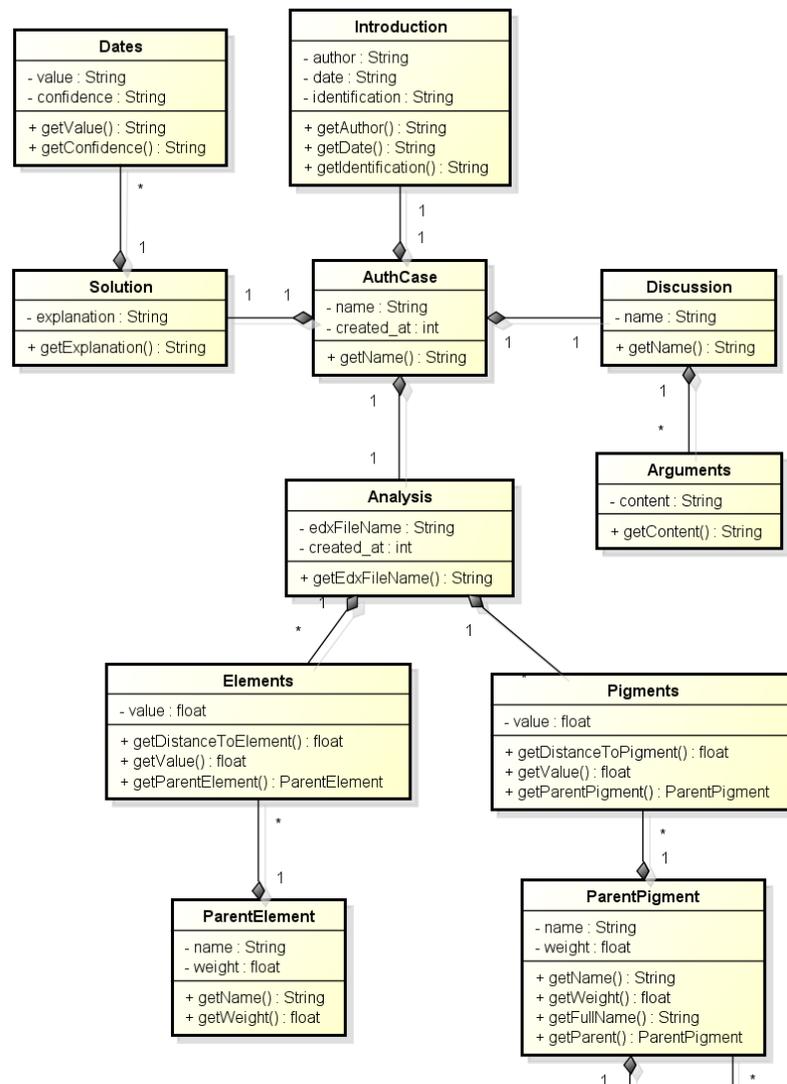
Figura 4.1 – Funcionamento do AuthenticPainting

No sistema, o usuário pode realizar uma consulta utilizando pigmentos, composições elementares, ou ambos. Neste caso, a recuperação de casos similares para cada um destes conjuntos de atributos de entrada acontece de maneira independente no sistema AuthenticPainting. Se o usuário decidir entrar com somente um tipo de dado para consulta (ou novo caso), seja ele pigmentos ou composições elementares, o sistema utilizará apenas a base de casos correspondente no processo de recuperação. Assim, apenas uma solução é apresentada para o usuário do sistema visto que a *Combinação 3* não será executada. Por consequência, a *Solução 3* não será obtida apresentada.

Do ponto de vista de implementação, as etapas de *Combinação 1* e *2* são equivalentes, visto que ambos os sistemas de CBR (baseados em pigmentos e baseados em composições elementares) geram resultados de recuperação no mesmo padrão de representação. Mas quando

as Soluções 1 e 2 podem ser geradas como resultados de consultas, o processo de Combinação 3 pode acontecer. Este processo é diferente dos processos anteriores de combinação, visto que devemos considerar que a solução gerada por meio de pigmentos é geralmente mais relevante no processo de autenticação de data de uma pintura. Logo, o sistema pode atribuir para resultados de recuperação baseados em pigmentos uma importância maior que para resultados de recuperação baseados em composições elementares.

Um caso de autenticação de pinturas é representado no sistema AuthenticPainting conforme o diagrama de classes da figura 4.2. As classes foram modeladas desta maneira com o objetivo de representar, o mais fielmente possível, a estrutura encontrada nos relatórios de autenticação de pinturas produzidos pelo especialista nesta aplicação.



powered by Astah

Figura 4.2 – Diagrama de Classes de um Caso de Autenticação do Sistema AuthenticPainting

O diagrama da figura 4.2 apresenta os principais componentes de um caso de autenticação de pinturas, onde a classe *AuthCase* representa um caso de autenticação. Um caso de autenticação tem um introdução, a qual é representada pela classe *Introduction*. Esta classe possui uma forte vínculo com a classe *AuthCase*. Logo, este relacionamento é representado por uma composição. Esta classe tem como objetivo armazenar as informações sobre o autor suposto da pintura, a data suposta e a identificação da mesma. Além disso, um caso de autenticação possui uma análise, a qual é representada no diagrama pela classe *Analysis*. A análise de um pintura contém os dados à respeito dos pigmentos históricos encontrados na pintura (classe *Pigments*) e das composições elementares obtidas em amostras de tinta da pintura (classe *Elements*). Tanto os diferentes tipos de pigmentos quanto os diferentes tipos de elementos químicos precisam estar previamente cadastrados no sistema. Deste modo, assegure-se que apenas informações válidas sejam inseridas em um caso de autenticação. Neste caso, as classes *ParentPigment* e *ParentElement* foram criadas para cumprir este requisito. Além disto, um determinado pigmento pode ser “filho” de outro pigmento, assim buscando modelar um relacionamento hierárquico entre pigmentos. Tendo em vista este comportamento, a classe *ParentPigment* possui um auto-relacionamento capaz de representar esta hierarquia no sistema. Todas as classes presentes na análise possuem um vínculo forte entre elas, onde estes relacionamentos são representados por composições. Em conjunto com a análise, um caso contém uma discussão de resultados de autenticação obtidos. A classe *Discussion* representa esta seção de discussão encontrada no processo de autenticação. Em uma discussão é possível identificar diferentes argumentos (observações, considerações, explicações, hipóteses, justificativas, etc), os quais são representados pela classe *Arguments*. Por fim, um caso de autenticação apresenta uma solução, a qual é representada pela classe *Solution*. Nesta solução, é possível identificar as justificativas alcançadas pelo especialista, as quais suportam os diferentes aspectos da sua decisão de autenticação. Em particular, em autenticação de data de pinturas, esta solução é composta por períodos de datas. Logo, a classe *Dates* é utilizada para representar estas datas obtidas na autenticação de pinturas históricas.

4.1. Exemplo de um caso de autenticação de pinturas: representação e codificação de um caso utilizando composições elementares

Em geral, o especialista retira várias amostras da pintura em diferentes partes do quadro (ver Apêndice A). Para cada amostra, diversas áreas são analisadas. Em cada área, são identificados os elementos químicos presentes (por exemplo: Al, Ti, etc) e quais as quantidades detectadas destes elementos. Para muitos casos, estas informações estão disponíveis no formato de uma planilha eletrônica. Nesta planilha, é possível obter os resultados de EDX obtidos nas análises no laboratório. Para importar esses resultados no sistema AuthenticPainting, os resultados de EDX destas planilhas devem ser armazenados em um arquivo texto separado por vírgulas (Comma-separated values – CSV). As figura 4.1.1 (A, B e C) apresenta as telas de inserção das composições elementares utilizando o arquivo CSV, onde o usuário deverá selecionar um arquivo CSV presente em seu computador.

O quadro 4.1.1 apresenta um exemplo de como os resultados EDX são armazenados em uma planilha eletrônica. Observe que normalmente várias amostras de uma pintura são retiradas. Cada uma destas amostras pode apresentar um número distinto de áreas. Os elementos químicos são identificados nestas áreas.

Para que seja possível utilizar estas composições elementares no sistema AuthenticPainting (neste caso, empregar essas composições no cálculo de similaridade entre casos de autenticação distintos), os resultados de EDX devem ser codificados. Neste processo, para cada amostra, considera-se uma única instância de cada elemento (por exemplo, independente de quantas amostras o elemento Al pode ser encontrado, apenas um elemento Al é considerado na representação de um caso). Dentre as várias ocorrências de um elemento químico em amostras retiradas da pintura, a representação deste elemento no sistema AuthenticPainting seleciona o maior valor de *WT (%)* do elemento. Maiores explicações e justificativas sobre este processo de representação de composições elementares podem ser encontradas em (SILVA, 2010).



Figura 4.1.1 (A) – Inserção de elementos com arquivo CSV

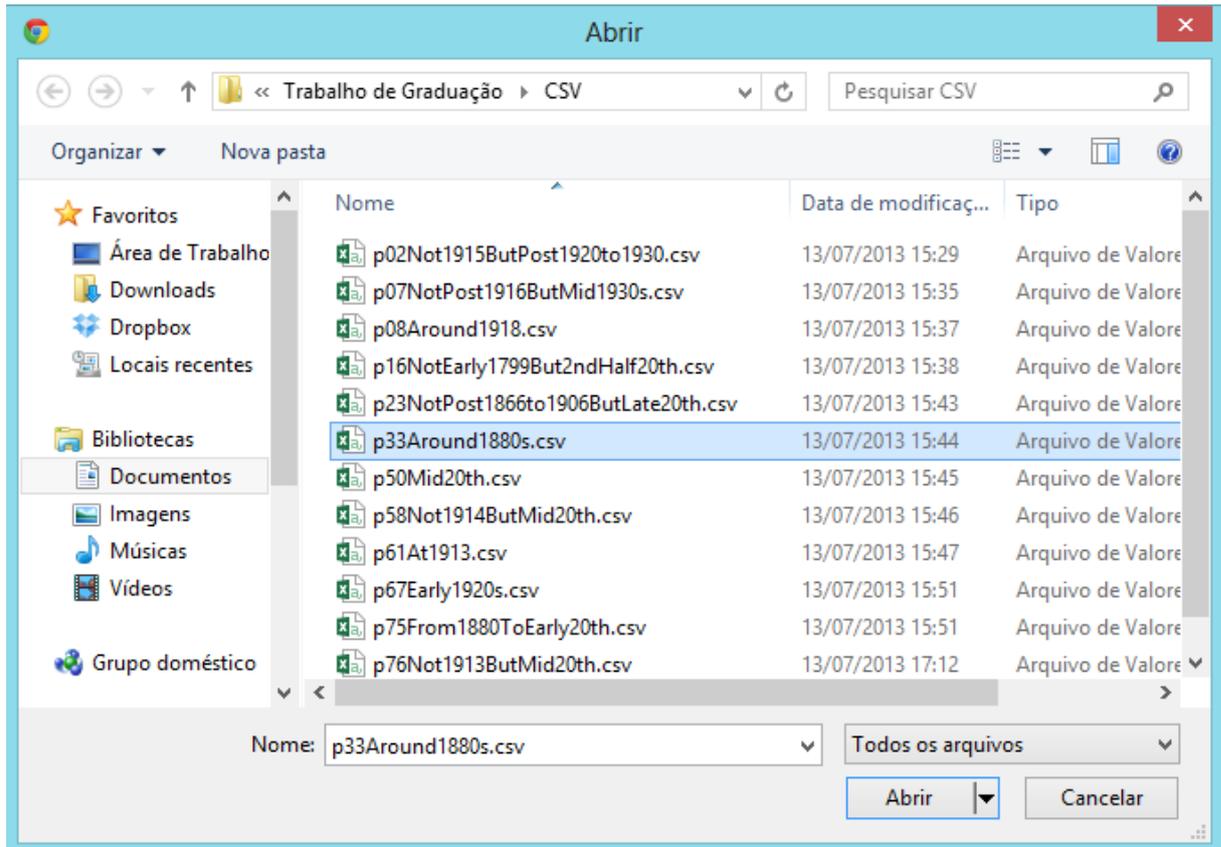


Figura 4.1.1 (B) – Caixa de seleção do arquivo CSV



Figura 4.1.1 (C) – Input após o arquivo selecionado

SAMPLE	AREA	ELEMENT	WT (%)
sample01	1	Al	23,5
sample01	2	Al	2,87
sample01	1	C	0,24
sample01	2	C	2,98
sample02	1	Al	3,67
sample02	2	Al	14,88
sample02	1	C	10,41
sample02	2	C	1,95
...

Quadro 4.1.1 – Exemplo de resultados EDX

Para o exemplo apresentando na quadro 4.1.1, resultado após o processo da interpretação é apresentado no quadro 4.1.2.

SAMPLE	AREA	ELEMENT	WT (%)
sample01	1	Al	23,5
sample01	2	C	2,98
sample02	2	Al	14,88
sample02	1	C	10,41
...

Quadro 4.1.2 – Resultados EDX após o primeiro passo de codificação

Após a etapa de seleção dos elementos de maior ocorrência em cada amostra, o sistema AuthenticPainting automaticamente codifica o valor de *WT (%)* para cada elemento utilizado os critérios demonstrados em (SILVA, 2010)

- Se $0 \leq WT(\%) \leq 1$, o elemento é codificado com o valor inteiro “1”. Neste caso, dizemos que existem traços do elemento (*trace*) na amostra em questão;
- Se $1 < WT(\%) \leq 10$, o elemento é codificado com o valor inteiro “2”. Neste caso, dizemos que existe pequena quantidade do elemento (*minor*) na amostra em questão;
- Se $WT(\%) > 10$, o elemento é codificado com o valor inteiro “3”. Neste caso, dizemos que existe grande quantidade do elemento (*major*) na amostra em questão.

Depois de efetuada esta primeira codificação, o sistema AuthenticPainting calcula a média aritmética de cada elemento em todas as amostras da pintura. Para o exemplo apresentado na quadro 4.1.2, temos como resultados os valores apresentados na quadro 4.1.3.

ELEMENT	VALUE
Al	3,0
C	2,5
...	...

Quadro 4.1.3 – Dados finais de elementos químicos do caso hipotético

Em resumo, estes valores médios de quantidades de ocorrências de composições elementares de amostras de pintura podem ser utilizados posteriormente no cálculo de similaridade entre dois casos. Para que isso seja utilizado, o usuário do sistema CBR deve

selecionar a composição elementar da pintura como propriedade relevante para o processo de consulta no sistema AuthenticPainting.

Cada um destes elementos químicos é previamente cadastrado no sistema AuthenticPainting. Esta base de elementos químicos representa todos os elementos que podem ser encontrados em uma amostra ou caso de autenticação. Além disto, cada elemento cadastrado no sistema tem um “peso” atribuído. Assim, cada um destes elementos pode ter uma relevância durante o processo de recuperação de casos de autenticação de pinturas da base de casos do sistema AuthenticPainting.

O sistema de CBR baseado em composições elementares foi desenvolvido levando em consideração a disponibilidade das planilhas resultantes das análises. Contudo, muitos casos de autenticação não disponibilizam estes arquivos. Logo, o sistema também permite que a composição elementar de uma pintura seja introduzida na base de casos manualmente. No modo de inserção manual, o especialista informa os elementos de maior quantidade para cada amostra. Não é necessário informar as diversas áreas de cada amostra, já que a inserção manual da composição elementar completa de um caso resultaria em uma tarefa muito complexa para o usuário. No modo de inserção manual, o sistema realiza apenas a última etapa de codificação demonstrada anteriormente, visto que o elemento de maior quantidade de cada amostra deve ser selecionado e inserido pelo usuário do sistema. A tela do modo de inserção manual é apresentada na figura 4.1.2.

Element Found	Value	Unit
Al	1.0	s1
Cd	0.24	s1
Cl	0.33	s1
Cu	0.81	s1
Pb	1.14	s1
Ti	2.69	s1
Zn	1.87	s1

Figura 4.1.2 – Modo de inserção manual de elementos

4.2. Processo de recuperação de casos utilizando composições elementares

Considere o seguinte cenário de utilização do sistema AuthenticPainting: a) uma base de casos composta de dois casos de autenticação de pinturas onde a solução do processo de autenticação desenvolvido para as pinturas é conhecida e, um caso alvo ou consulta o qual se deseja autenticar a suposta data de criação da pintura. Neste cenário exemplo, a base de casos é composta pelos casos **PaintingA** e **PaintingB**. Para ambos os casos, já foi possível identificar o período de datas que as pinturas pertencem. Além disto, dados sobre composições elementares e pigmentos estão disponíveis para estes casos na base de casos do sistema.

Usando o sistema AuthenticPainting, um especialista deseja autenticar a data da pintura **PaintingC** através da utilização de valores de composições elementares obtidos nesta pintura C usada como consulta. Considere o seguinte fragmento do relatório da pintura **PaintingC**.

*This report presents the results of a scientific examination of a painting referred to here as the **PaintingC**, thought to be by **Author of PaintingC**. The painting measures ... but is apparently interpreted as 1949. Other painting by **Author of PaintingC** with comparable stylistic elements can be found in the period 1947-54, providing a date range for comparison.*

Figura 4.2.1 – Introdução PaintingC

Observe que o caso PaintingC ainda não apresenta informações sobre pigmentos. Este cenário onde pigmentos ainda não foram identificados é encontrado normalmente no início do processo de autenticação de uma pintura. Desta forma, o especialista pode utilizar os dados sobre composições elementares, os quais já estão disponíveis, para obter uma “possível solução” para o processo de autenticação (uma solução alcançada rapidamente no processo de autenticação). Outro fator relevante de observarmos é que, ao recuperar casos similares utilizando composições elementares, o especialista pode identificar (nas informações presentes na seção “análise” dos casos recuperados) os possíveis pigmentos que ele poderá encontrar em uma etapa posterior da autenticação da pintura. Após a identificação, estes pigmentos podem vir a ser utilizados como entrada na recuperação de casos similares em relação a presença de pigmentos. Esta recuperação de casos utilizando informação sobre pigmentos é descrita na seção 4.4 deste TG.

Para a pintura **PaintingC**, a aplicação da técnica EDX gerou os resultados apresentados na quadro 4.2.1. Esta tabela apresenta os resultados após o processo de codificação descrito na seção 4.1.

Al	1,44
----	------

As	0,55
Ba	1,55
Br	0
Ca	1,55
Cd	0,77
...	...

Quadro 4.2.1 – Codificação dos resultados EDX para PaintingC

Após a entrada dos dados da Tabela X no sistema AuthenticPainting, os dois casos da base de casos (*PaintingA* e *PaintingB*) são recuperados como resultados de uma consulta, juntamente com a similaridade de cada um deles em relação ao caso consulta **PaintingC**. A figura 4.2.2 (A, B, C e D) apresenta fragmentos da pintura *PaintingA*.

(A) *This report presents the results of a scientific examination of a painting referred here as the **PaintingA**. ...The painting is unsigned and undated but presents itself as a work of a type produced by **Author of PaintingA** around 1915-1916.*

(B) *The following pigments were identified: zinc oxide ('zinc white'); cobalt aluminium oxide ('cobalt blue'); mercury(II) sulfide ('vermilion'), as a so-called 'dry-process' type; zinc potassium chromate ('zinc chromate'). A carbon-based black...*

(C) *As reported above, no materials known to be anachronistic to the period of **Author of PaintingA** works were found. For example, pigments of practical introduction later than 1920s such as titanium dioxide white variants and ... We can however make some more direct comparisons to **Author of PaintingA** known practices. There are also in accord with pigments identified in works of roughly the same period by **Author of PaintingA**.*

(D) *We may summarise the findings as follows: The pigments found in the painting are consistent with a supposed date for the painting of 1915-16, though they have been available across a much wider range than this. Where multiple types of a pigment are known to exist...*

Figura 4.2.2 – Conclusões *PaintingA*

O caso **PaintingA** apresenta os resultados EDX codificados conforme quadro 4.2.2.

Al	0,42
----	------

As	0
Ba	0
Br	0
Ca	0,57
Cd	0
...	...

Quadro 4.2.2 – Resultados EDX codificados para a pintura PaintingA

O caso **PaintingB** possui as informações apresentadas pelos fragmentos do caso observados na figura 4.2.3 (A, B, C e D).

*(A) This report presents the results of a scientific examination of a painting referred to here as **PaintingB**. ...The Painting is unsigned and undated but thought to be by **Author of PaintingB** around 1906/7...*

*(B) The following pigments were identified: zinc oxide ('zinc white'); synthetic ultramarine; mercury(II) sulfide ('vermillion'), as a so-called 'wet-process' type; chromium oxide hydroxide ('viridian'); cadmium sulfide ('cadmium yellow');
....*

(C) First, the primary paint matrix is throughout based on zinc oxide. The occurrence of a zinc white in the painting is important for the probable objective dating of the work since there is a broad evolution in the use of different white hiding pigments across the period...

*(D) This analysis has confirmed the presence of pigments within the primary structure of the painting – notably a zinc white ...There was also good correspondence with the range of pigments known to have been used by **Author of PaintingB**. It therefore appears reasonable to conclude that the material structure of the painting is consistent with what we would expect of a painting by **Author of PaintingB** from the supposed period around 1906/07.*

Figura 4.2.3 – Conclusão PaintingB

Os resultados de EDX da **PaintingB** são apresentados na quadro 4.2.3 a seguir.

Al	1,08
As	0
Ba	1,16
Br	0
Ca	1,75

Cd	1,16
...	...

Quadro 4.2.3 – Resultados EDX codificados para a pintura PaintingB

Utilizando o cálculo de similaridade apresentado na seção 2.1.2, é possível obter a similaridade de cada um dos casos da base de dados com o caso alvo **PaintingC**. Esta similaridade é apresentada na quadro 4.2.4.

Caso	Similaridade
PaintingB	88,33%
PaintingA	82,72%

Quadro 4.2.4 – Baseado em composições elementares, casos similares a PaintingC

A partir destes resultados, é possível observar que **PaintingB** apresenta a maior similaridade com **PaintingC**. Desta forma, este seria o caso com a “melhor” solução a ser reutilizada como solução para **PaintingC**. Entretanto, o processo de reutilização de soluções passadas também pode buscar a reutilização das soluções encontradas nos demais casos mais similares recuperados. A partir dos resultados obtidos nesta recuperação de casos, a combinação das soluções dos casos **PaintingA** e **PaintingB** pode vir a fornecer uma solução mais completa para o novo problema. A figura 4.2.4 apresenta a tela de recuperação de casos similares ao caso **PaintingC**.

A figura 4.2.4 permite visualizar os intervalos de datas que o sistema acredita que sejam prováveis para a pintura usada como consulta. Assim, é possível que o especialista analise as possíveis soluções para o caso de entrada, contribuindo para o desenvolvimento do processo de autenticação de data da pintura.

Além dos intervalos de datas apresentados, o gráfico descrito na figura 4.2.4 informa a importância de cada uma das datas no eixo vertical. Desta forma, é possível apresentar resultados que o especialista (ou mesmo o sistema) acredita que sejam mais prováveis para solucionar o caso usado como consulta. No exemplo da figura 4.2.4, podemos observar que o sistema retornou 2 casos da base de casos. Assim como descrito nestes casos, eles apresentam uma importância padrão para cada uma das datas atribuídas para suas pinturas, visto que nenhuma destas datas foi especificada pelo especialista com um modificador de importância tal como *Maybe* ou *Likely*.

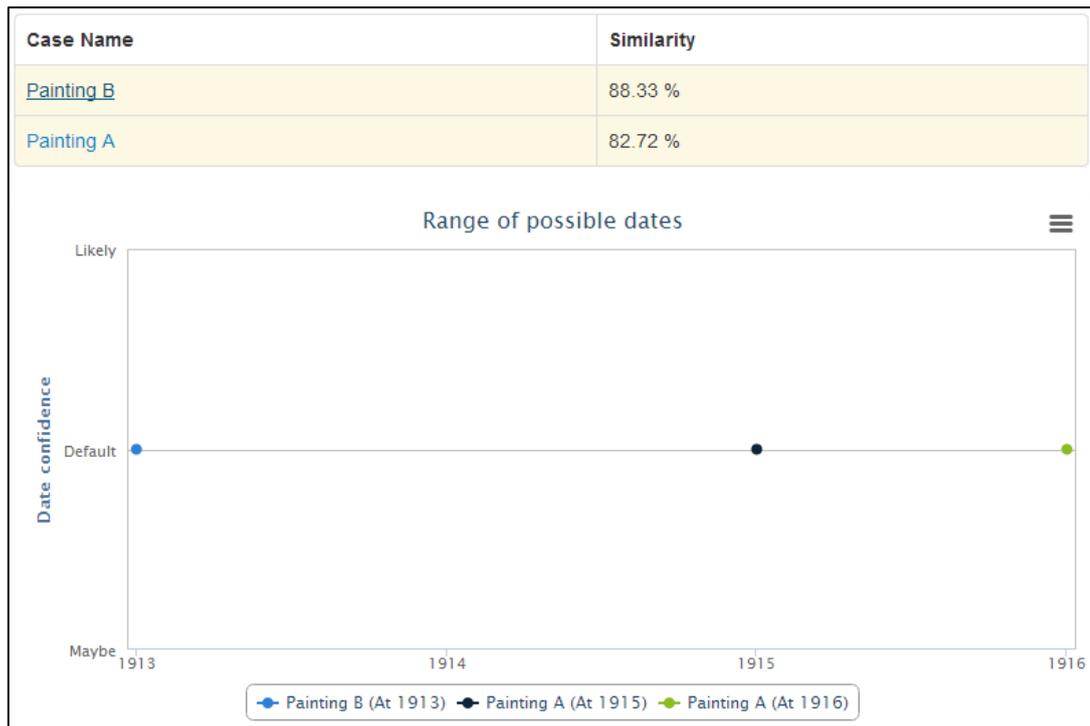


Figura 4.2.4 – Resultado da recuperação utilizando composição elementares

4.3. Exemplo de caso: representação de um caso utilizando pigmentos

Neste TG, os pigmentos presentes em uma pintura são utilizados na construção do sistema de CBR para autenticação de pinturas. Como já visto no Capítulo 3, os pigmentos são considerados mais relevantes para determinar a data de uma pintura (ou seja, eles possuem maior informação sobre data do que composições elementares).

A partir do estudo dos relatórios de autenticação, é possível observar que o especialista pode descrever um mesmo pigmento de diversas formas. Mas para a utilização de pigmentos no AuthenticPainting, foi necessário padronizar o método de representação destes. No sistema, todos os pigmentos possíveis de serem utilizados em um caso de autenticação estão representados de forma hierárquica. Esta hierarquia é representada em um arquivo no formato XML (*Extensible Markup Language*). A figura 4.3.1 apresenta um trecho deste arquivo XML.

Este arquivo é carregado e processado no início da execução do sistema. Desta forma, cada pigmento é armazenado no banco de dados do sistema. Esta base de pigmentos pode ser expandida (ou mesmo ajustada) através da interface administrativa do sistema. A figura 4.3.2 apresenta uma tela desta interface onde é possível gerenciar a base de pigmentos do sistema.

A inserção de um novo caso no sistema, o qual possua pigmentos, ocorre de modo semelhante à inserção manual de composições elementares. Neste caso, o usuário deve informar manualmente os pigmentos identificados na pintura em questão. A figura 4.3.3 apresenta a tela de aquisição de um caso baseado em pigmentos.

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <pigments-table>
3    <pigment>
4      <name>Root</name>
5      <weight>0.0</weight>
6      <pigments>
7        <pigment>
8          <name>arsenic</name>
9          <weight>1.0</weight>
10         <pigments>
11           <pigment>
12             <name>orpiment-realgar</name>
13             <weight>1.0</weight>
14           </pigment>
15         </pigments>
16       </pigment>
17     <pigment>
18       <name>barium</name>
19       <weight>5.0</weight>
20       <pigments>
21         <pigment>
22           <name>barium-sulfate</name>
23           <weight>5.0</weight>
24         </pigment>
25       </pigments>
26     </pigment>

```

Figura 4.3.1 – Base de Pigmentos em arquivo XML

Figura 4.3.2 – Interface administrativa: pigmentos

Na figura 4.3.3, é possível notar que o usuário possui alternativas de descrição de pigmentos em diferentes níveis de detalhe (e granularidade). Durante o processo de autenticação, o especialista pode encontrar evidências da presença do pigmento *cadmium* em uma pintura. Porém, este especialista pode não obter mais detalhes a respeito deste pigmento, o que impossibilita que o pigmento seja identificado como *cadmium-sulfide* ou *cadmium-selenide-sulfide*, por exemplo. Neste caso, o especialista apenas informa ao sistema que a pintura possui *cadmium*. Não existe a necessidade de informar qual das prováveis variantes o pigmento pertence, os quais podem não ser possíveis de ser conhecidas a partir das análises realizadas na pintura.

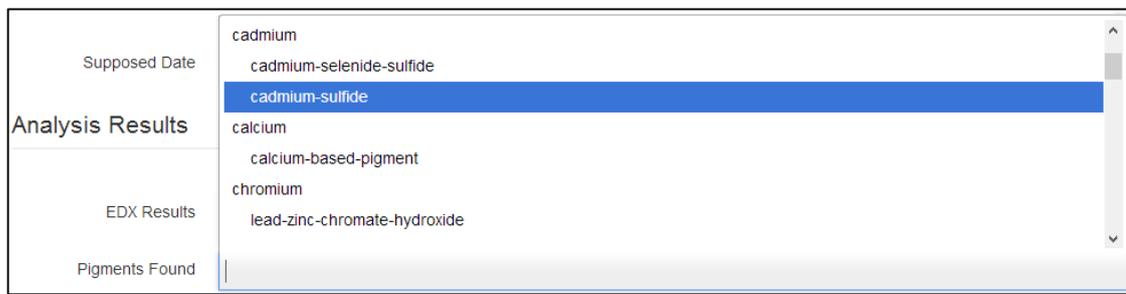


Figura 4.3.3 – Inserção de pigmentos na aquisição de um caso

O processo de representação dos pigmentos descrito neste TG pode permitir que o sistema obtenha um melhor resultado no cálculo da distância entre dois pigmentos. Neste caso, é possível obter um valor de distância local entre *cadmium* e *cadmium-selenide-sulfide* por exemplo. Isto se deve a representação hierárquica dos pigmentos. Esta forma de representação busca expandir o método descrito no trabalho de (SILVA, 2010), onde o sistema iria considerar a presença de dois pigmentos distintos. Neste caso, a distância local entre eles seria máxima. Porém, sabemos que entre *cadmium* e *cadmium-selenide-sulfide* existe algum grau de similaridade, assim como para outros tipos de pigmentos implementados no sistema.

4.4. Processo de recuperação de casos utilizando pigmentos

No processo de recuperação de casos baseados em pigmentos, podemos descrever um exemplo de recuperação onde a base de casos é composta pelos casos **PaintingA** e **PaintingB**. Neste exemplo, estes casos apresentam dados sobre os pigmentos. Considere os seguintes

pigmentos para a pintura **PaintingC**: *synthetic ultramarine, cobalt blue, zinc oxide*, dentre outros. Observe que a Pintura C é um caso alvo, ou caso que o usuário deseja autenticar a suposta data de criação da pintura.

Para realizar a consulta dos casos similares a pintura **PaintingC**, o usuário do sistema deve informar estes pigmentos no formulário de consulta do sistema. A inserção dos pigmentos é realizada da maneira apresentada pela figura 4.3.3. Após o usuário do sistema informar os pigmentos da pintura **PaintingC** como entrada na consulta, que é agora baseada em similaridades de pigmentos encontrados em diferentes casos de autenticação, o sistema AuthenticPainting recupera os casos **PaintingA** e **PaintingB** juntamente com suas respectivas similaridades ao caso de entrada.

Quando pigmentos são utilizados no processo de recuperação de casos, o sistema busca, em cada caso da base de casos, pela existência do pigmento em questão. O cálculo da distância ocorre da seguinte forma:

- Se ambos os casos possuírem o pigmento, dizemos que a distância local para este pigmento é “zero” (distância mínima).
- O cálculo de similaridade considera os pigmentos do caso de consulta que não estão presentes no caso da base de casos, onde para cada um destes pigmentos a distância local é definida como “um” (distância máxima).
- Para cada caso da base de casos é necessário considerar os pigmentos que não estão presente no caso de consulta e, para estes pigmentos, dizemos que a distância local “um”.
- O cálculo de similaridade não considera os pigmentos que não estão presentes tanto no caso de consulta quanto no caso da base de casos em questão.

Uma vez que sejam efetuados os cálculos das distâncias locais para cada um dos casos da base de casos do sistema, o cálculo da distância Euclidiana (ou distância global entre casos) é realizado. Para este cenário, o sistema pode gerar resultados tais como os apresentados na figura 4.4.1.

Conforme pode ser observado na figura 4.4.1, as consultas baseadas em pigmentos são respondidas no mesmo formato de representação utilizado para consultas baseadas em composições elementares. Entretanto, assim como demonstrado em (SILVA, 2010), espera-se que os resultados baseados em pigmentos gerados pelo sistema de CBR sejam mais confiáveis

que aqueles gerados pelo sistema de CBR usando composições elementares (visto que pigmentos carregam mais informações sobre a data para a autenticação de pinturas).

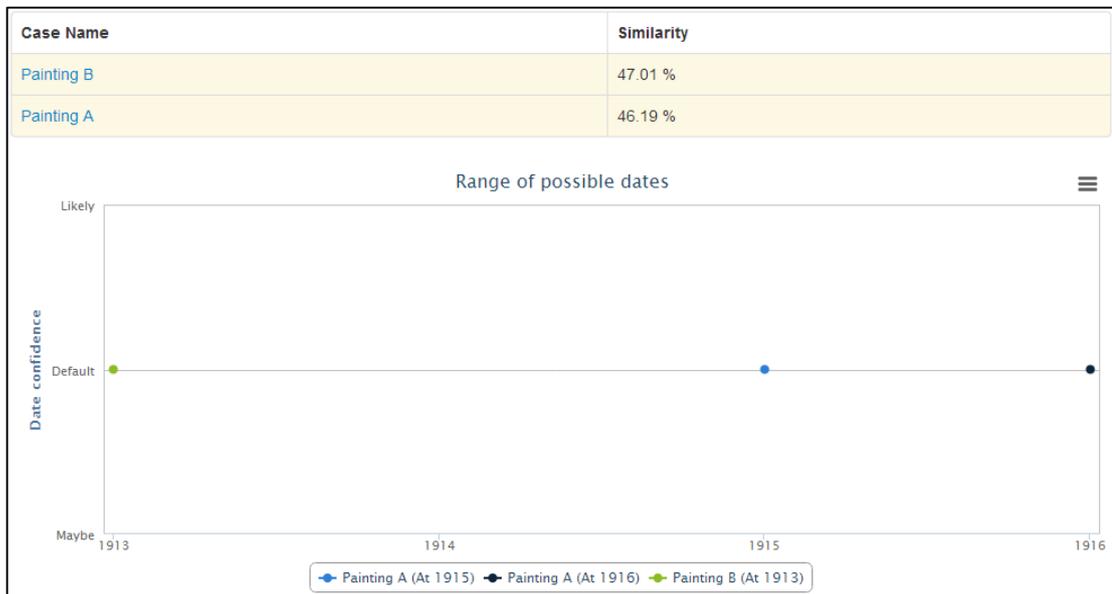


Figura 4.4.1 – Recuperação com de casos similares utilizando pigmentos

4.5. Problema de Combinação de Datas em Autenticação de Pinturas

As datas propostas para pinturas encontradas nos relatórios de autenticação são expressas de diversas maneiras. Tais descrições são apresentadas em um formato textual, sem uma estrutura padronizada. Logo, para que seja possível a combinação destas datas no AuthenticPainting (ou seja, a combinação de resultados de recuperação de casos similares para uma consulta), um padrão para a representação destas datas deve ser definido.

Em geral, as data estabelecidas nestas conclusões de autenticação não se referem à uma data específica. Muitas vezes, estas datas são representadas por um período de tempo, tal como uma década (no caso de pinturas mais recentes) ou um século (no caso de pinturas mais antigas) Por exemplo, Post 1920s ou Mid 17th. Além disto, em alguns relatórios de autenticação, pinturas podem ser atribuídas a mais de uma possibilidade de data, assim refletindo as diferentes evidências para data encontradas no processo de análise da pintura. Em alguns casos, estas várias datas atribuídas a uma pintura possuem uma “precisão” associada. Muitas vezes, tal precisão é expressa pelo especialista por termos *likely* e *maybe*, por exemplo. Quando uma pintura possui diversas datas supostas, estes termos podem ser usados para designar qual das

datas é a data mais provável da pintura. Este é um julgamento baseado na combinação das evidências relevantes para atribuição da data encontradas na pintura.

Neste TG, para propor uma representação para datas obtidas nestes processos de autenticação, uma análise de como o especialista costuma descrever tais datas foi realizada. A ideia principal foi identificar uma forma de representação que pudesse capturar (o mais fielmente possível) o maior número possível de datas contidas em relatórios de autenticação. Além disso, o objetivo foi desenvolver um padrão de representação que pudesse ser computável, onde a adoção do padrão pudesse ser relativamente simples para o usuário do sistema nesta aplicação. Em conclusão, a representação de datas proposta neste TG possui os seguintes elementos:

- *At data*: representa uma data ou período específico. Este padrão é utilizado quando a pintura possui esta data ou período já atribuídos (tal como uma data escrita no quadro) e a autenticação identifica que não existem evidências que sejam contrárias a esta atribuição (*At 1915*, por exemplo)
- *Post data*: representa um período que seja após uma determinada data. Em geral, este padrão é empregado quando o especialista encontra evidências em uma pintura que impossibilitem a mesma de pertencer a um período anterior a uma determinada data. Por exemplo, se for identificado algum tipo particular de pigmento sintético na pintura, a data é *Post 1945*, já que estes pigmentos só foram inventados durante a Segunda Guerra Mundial.
- *Around data*: este padrão de data representa um período menos específico. Em geral, este padrão é utilizado quando não são identificadas evidências significativas para atribuição de uma data mais específica. Neste caso, por exemplo, não são identificados pigmentos que possuam uma data clara onde a sua utilização é encontrada. Porém, tais pigmentos estão disponíveis na data proposta para a pintura (*Around 1950*, por exemplo). Este padrão também pode ser usado em pinturas bastante antigas, visto que os pigmentos encontrados nestas pinturas foram, em geral, utilizados por extensos períodos de tempos, impossibilitando a determinação de uma data mais precisa para a pintura.
- *From data To data*: este padrão é importante quando são identificadas evidências (normalmente pigmentos) que possuem uma data/período de introdução (ou seja, que os pigmentos começaram a aparecer em pinturas) e uma data/período que foram

abandonadas (por exemplo, que os pigmentos foram substituídos pelos fabricantes por formulações mais modernas: *From 1920s to 1950s*, por exemplo).

4.6. Processo de combinação dos resultados gerados por pigmentos com os resultados gerados por composições elementares

Após a recuperação de casos similares utilizando pigmentos e composições elementares, o sistema é capaz de combinar estes resultados. Neste caso, o sistema pode atribuir uma maior importância para os resultados baseados em pigmentos gerados pelo sistema de CBR. Neste caso, o usuário tem a liberdade de configurar os pesos de acordo com a sua necessidade de análise durante o processo de autenticação de novas pinturas. Para realizar a combinação dos resultados, o sistema utiliza a equação 3 descrita abaixo.

$$sim(a, b)_{comb} = \frac{(sim(a,b)_{pig} * weight_{pig}) + (sim(a,b)_{elem} * weight_{elem})}{weight_{pig} + weight_{elem}} \quad (3)$$

Onde:

- $sim(a, b)_{comb}$ representa a similaridade entre os casos “a” e “b” obtida como resultado do processo de combinação de pigmentos e composições elementares;
- $sim(a, b)_{pig}$ representa a similaridade entre os casos “a” e “b” obtida através do sistema de CBR baseado em pigmentos apenas;
- $sim(a, b)_{elem}$ representa a similaridade entre os casos “a” e “b” obtida através do sistema de CBR baseado em composições elementares apenas;
- $weight_{pig}$ representa o peso atribuído para a similaridade entre os casos “a” e “b” obtida através do sistema de CBR baseado em pigmentos.
- $weight_{elem}$ representa o peso atribuído para a similaridade entre os casos “a” e “b” obtida através do sistema de CBR baseado em composições elementares.

Assim como configurado no sistema AuthenticPainting, este cálculo é efetuado para cada um dos quatro resultados com similaridade maior que 50% de cada sistema. Nos testes realizados com este processo de combinação, o sistema foi configurado para que os pigmentos tivessem um peso de três vezes maior que o peso de resultados obtidos por meio de composições elementares. Após a combinação, os casos são ordenados pela similaridade resultantes e os

quatro casos mais similares retornados são exibidos como o resultado do processo de combinação.

O objetivo deste método de combinação é fornecer um meio de confirmar os resultados baseados em pigmentos, visto que estes são mais importantes para o processo de autenticação das pinturas.

Ao utilizarmos este método para os exemplos apresentados na seção 4.2 (composições elementares) 4.4 (pigmentos), o sistema foi capaz de fornecer um resultado que possibilita o especialista confirmar os resultados gerados pelo sistema baseado em pigmentos. A figura 4.6.1 apresenta o resultado da combinação e a afirmação da pintura **PaintingB**, como a pintura mais similar ao caso alvo **PaintingC**.

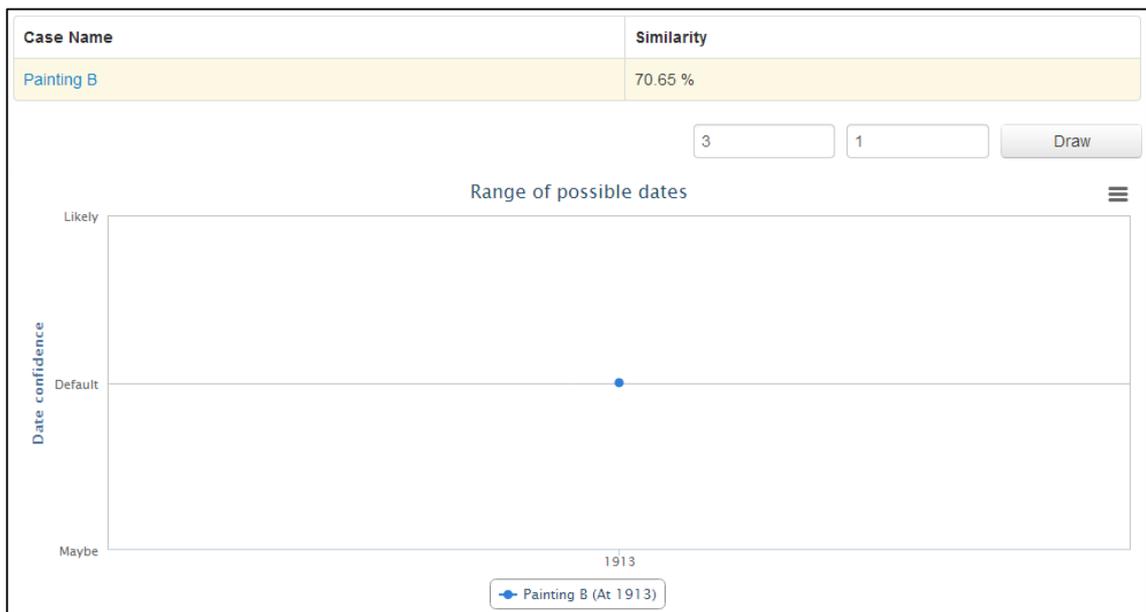


Figura 4.6.1 – Confirmação da pintura **PaintingB** como a mais relevante para o processo de autenticação da pintura alvo **PaintingC**

5. TESTES DO SISTEMA

Para a avaliação do sistema, 12 casos contidos na base de casos do sistema foram selecionados para serem usados como consultas. Estes casos selecionados incluem pinturas autenticadas como reais e pinturas consideradas como falsas. Todos os casos de teste possuem dados sobre ambas composições elementares e pigmentos. Além disto, estes casos escolhidos foram atribuídos para diferentes períodos de tempo. Deste modo, espera-se testar o sistema no suporte a autenticação de datas de pinturas antigas e de pinturas mais recentes (embora a base de casos do sistema possua mais casos de autenticação de pinturas mais novas, ou pinturas do século 20, conforme pode ser observado na Figura 5.1 – adaptada de (SILVA, 2010)).

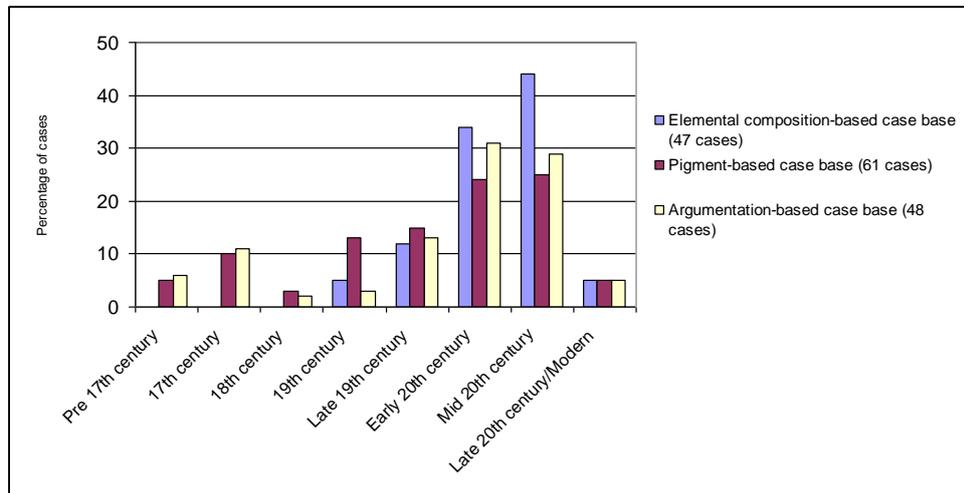


Figura 5.1 – Percentagem de casos em diferentes períodos de tempo presentes nas bases de casos

Durante a execução dos testes, o *leave one out and test* (WATSON, 2003) foi a estratégia utilizada. Neste caso, o caso a ser testado foi retirado da base de casos do sistema e utilizado como consulta. No formulário de consulta, estes casos usados como teste foram inseridos através das suas composições elementares e dos seus pigmentos. O sistema então busca pelos quatro casos mais similares para cada um destes dados de entrada (pigmentos e composições elementares). Assim como configurado nos testes do sistema, os casos retornados como respostas para estas consultas devem possuir uma similaridade com o caso de teste de no mínimo 50%.

Após as recuperações de casos similares via composições elementares e pigmentos, o sistema realiza a combinação destes resultados obtidos para a consulta alvo. Para isso, uma importância três vezes maior foi atribuída para resultados obtidos via pigmentos, enquanto a importância dos resultados obtidos via composições elementares permaneceu inalterada. Com a atribuição de diferentes pesos para os pigmentos, é possível avaliar a relevância de combinarmos diferentes tipos de dados a fim de obter uma possível melhoria das soluções obtidas no sistema de CBR.

Ao final de cada teste, uma análise crítica da utilidade dos três diferentes tipos de resultados de recuperação (pigmentos, composições elementares, ou ambos combinados onde pigmentos têm 3 vezes mais importância que composições elementares) foi realizada. Esta análise levou em consideração os casos retornados e a data real da pintura usadas como consulta ou teste. Por fim, a análise considera as seguintes opções:

- Bastante útil: o sistema recuperou casos com grande similaridade, onde as datas obtidas nos casos recuperados estão concentradas em curto intervalo de tempo;
- Útil: o sistema recuperou casos, onde as datas destes casos estão próximas as datas aceitáveis para a pintura usada na consulta;
- Talvez útil: o sistema recuperou apenas alguns casos como resposta para a consulta, onde as datas destes casos são aceitáveis para a pintura alvo. Esta avaliação também pode ser utilizada quando os casos recuperados possuem uma similaridade baixa com o caso usado na consulta.
- Não útil: o sistema recuperou casos em que as datas contidas nos casos recuperados não são aceitáveis para a pintura usada na consulta.
- Não aplicável: esta avaliação pode ser usada quando o caso usado na consulta possui complicações normalmente associadas a falsificações, as quais o sistema não possui conhecimento para acertar a solução do problema. Além disso, esta avaliação também é relevante quando os casos da base de casos não possuem data similares a data que está sendo autenticada na pintura usada como consulta (ver datas das pinturas armazenadas como casos no sistema na Figura 5.1) Por fim, esta opção permite indicar que o sistema não recuperou casos via pigmentos que pudessem ser comprovados pela recuperação baseada em composições elementares.

É importante salientar que, em autenticação de pinturas, o conceito de “datas aceitáveis” é bastante relativo. Se considerarmos uma pintura mais antiga (anterior a metade do século 19, por exemplo), as “datas próximas” para esta pintura abrangem um intervalo consideravelmente maior que o intervalo considerado aceitável para pinturas mais recentes. Por exemplo, um intervalo de 50 anos ou mais é considerado “aceitável” para autenticações de pinturas antigas. Logo, a análise da utilidade dos resultados obtidos nos testes do sistema AuthenticPainting leva em consideração estes fatos relacionados a análise de datas próximas ou aceitáveis.

5.1. Um exemplo de teste do sistema utilizando a pintura p75From1880ToEarly20th como consulta

Para a realização deste teste, a pintura p75 foi retirada da base de casos e os seus dados sobre composições elementares e pigmentos foram inseridos no formulário de consulta do sistema. Para esta pintura, acredita-se que sua data de criação tenha sido de 1880 até início do século 20. Utilizando dados de pigmentos no processo de recuperação, o sistema gerou os resultados apresentados na figura 5.1.1.

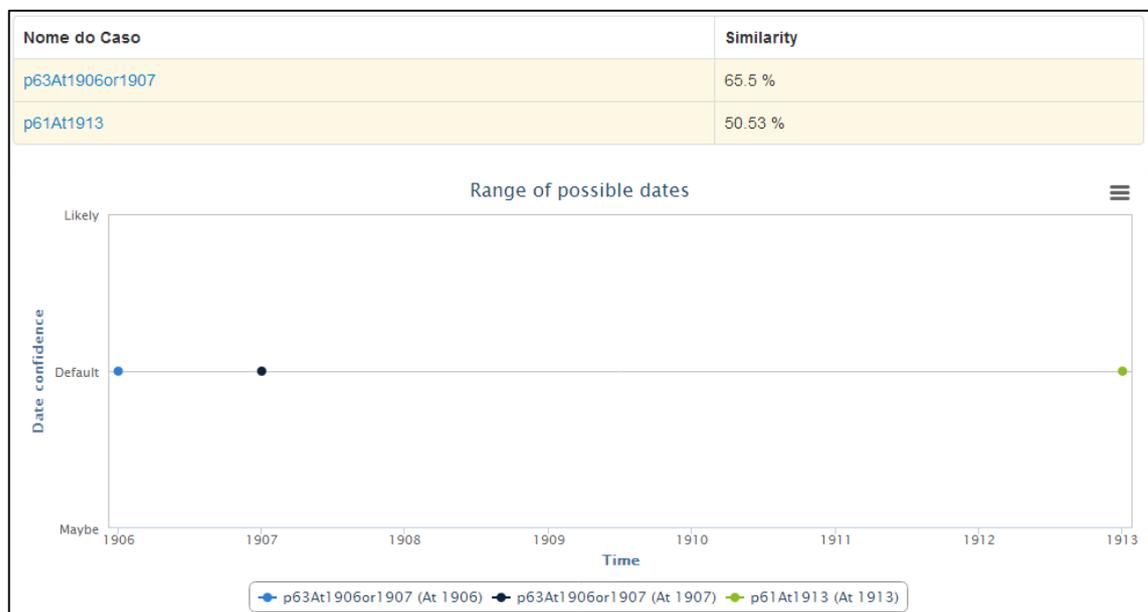


Figura 5.1.1 – Resultados da consulta da pintura p75 utilizando pigmentos

Como é possível visualizar na figura 5.1.1, o sistema retornou dois casos com similaridade superior a 50%. Ambos os casos possuem um foco em datas próximas às datas

autenticadas da pintura de teste p75. Deste modo, é possível afirmar que o resultado gerado é bastante útil para auxiliar o processo de autenticação da pintura. O quadro 5.1.1 apresenta a análise da utilidade dos resultados obtidos.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p75.				

Quadro 5.1.1 – Análise dos resultados da pintura p75 para pigmentos

Para esta pintura p75, os resultados da recuperação dos casos similares utilizando composições elementares são apresentados na figura 5.1.2.

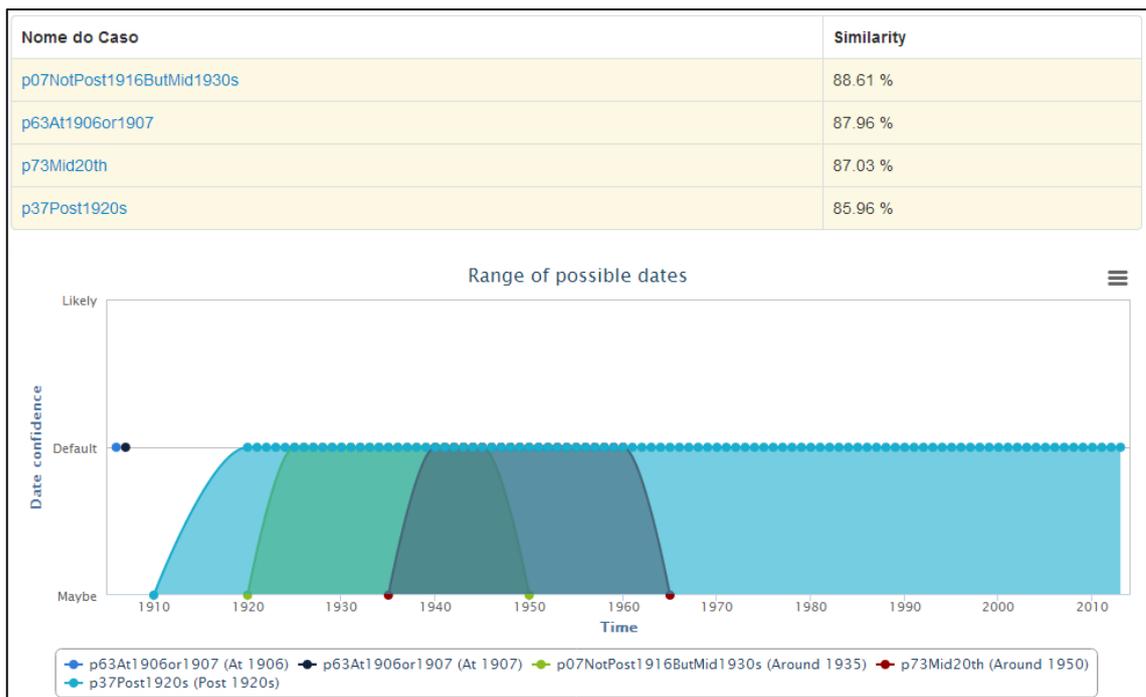


Figura 5.1.2 – Resultados da consulta da pintura p75 utilizando composições elementares

Usando composições elementares, o sistema retornou um maior número de casos similares e com similaridades maiores. Além disto, todos os casos retornados apresentam datas próximas às datas autenticadas da pintura de teste, embora os intervalos de dados retornados

não focalize datas do início do século 20, assim como autenticado na pintura p75 usada como teste no sistema.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p75.				

Quadro 5.1.2 – Análise dos resultados da pintura p75 para composições elementares

Por fim, resultados combinando os casos retornados pelo sistema de CBR baseado em pigmentos e o sistema de CBR baseado em composições elementares foram gerados. Neste teste, consideramos os resultados baseados em pigmentos como três vezes mais relevantes que os resultados baseados em composições elementares. A figura 5.1.3 apresenta os resultados obtidos.

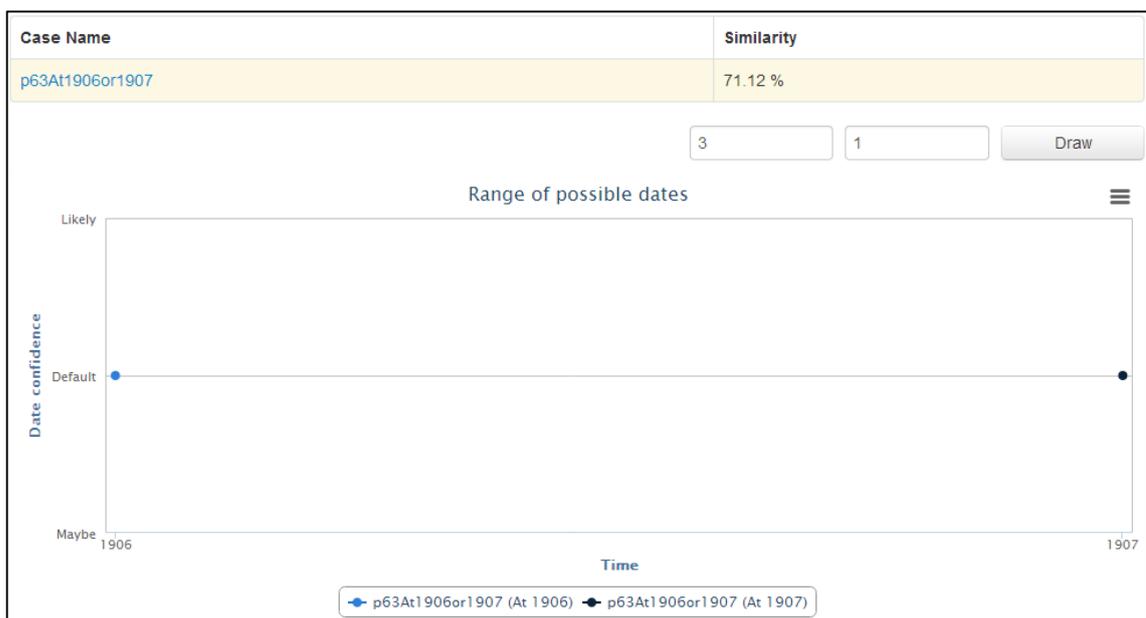


Figura 5.1.3 – Resultado da combinação utilizando peso 3 para os pigmentos

Neste teste, o mesmo caso p63 foi retornado. Porém, a similaridade com o caso teste reduziu visto que ao atribuir um peso maior para os pigmentos, o resultado da combinação se

aproximou dos resultados obtidos quando somente os pigmentos foram utilizados. Os resultados da análise são apresentados no quadro 5.1.4.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
A pintura recuperada apresenta datas próximas as datas reais da pintura teste p75.				

Quadro 5.1.3 – Análise dos resultados da pintura p75 utilizando peso 3 para os pigmentos

Tendo em vista os resultados obtidos para a pintura p75, foram gerados resultados que, em uma situação real de autenticação, poderiam auxiliar o especialista durante o processo de análise da pintura alvo. Assim como exemplificado para a pintura p75, testes com mais 11 pinturas foram realizados, onde todos seguiram esta mesma metodologia. Para os demais testes executados, é importante salientar que, quando diferentes pinturas aparecem nos resultados de recuperação baseados em pigmentos e baseados em composições elementares, o método de combinação explorado não consegue recuperar casos que possuam similaridade superior a 50% (ou seja, o resultado de recuperação é vazio no sistema). Os resultados dos testes, juntamente com a análise e a justificativa das soluções encontradas estão disponíveis no Anexo A deste TG.

Após a realização de todos os testes e a análise crítica dos resultados obtidos, é possível apresentar uma análise da eficiência do sistema. Esta análise é construída de acordo com a utilidade dos resultados no auxílio da obtenção das soluções para os casos de autenticação de data usados como consultas no sistema AuthenticPainting.

A avaliação dos resultados apresentados nas figuras 5.1.4, 5.1.5 e 5.1.6 permite concluir que o método de combinação proposto tem a necessidade de que ambas as recuperações retornem casos em comum. Caso contrário, o método de combinação testado não é usável, mesmo considerando que os resultados obtidos quando consultas independentes são executadas (baseadas em pigmentos e baseadas em composições elementares, mas não combinadas segundo o método de combinação proposto) são, de certa forma, úteis ao processo de autenticação de data.

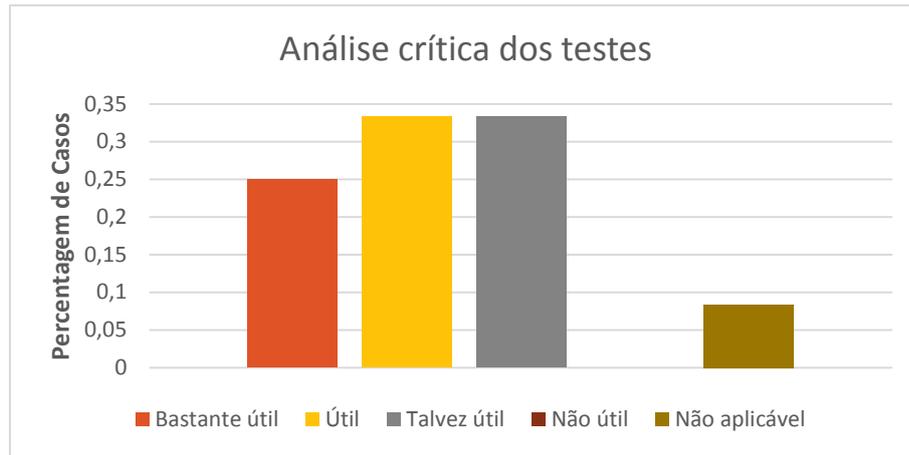


Figura 5.1.4 – Resultado das análises das recuperações por pigmentos

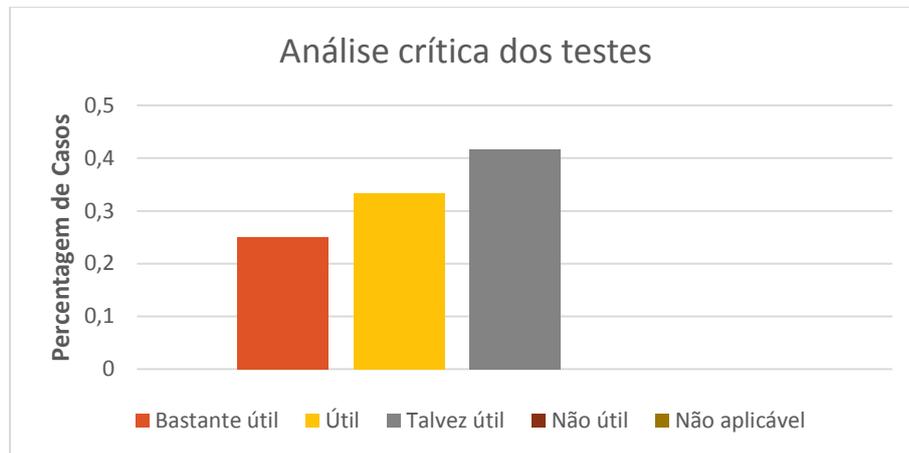


Figura 5.1.5 – Resultado das análises das recuperações por composições elementares

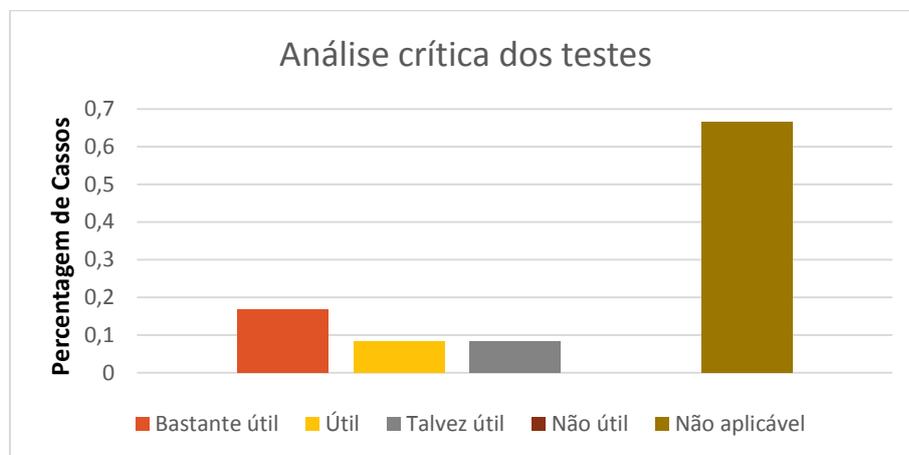


Figura 5.1.6 – Resultado das análises com peso 3 para os pigmentos

6. CONCLUSÕES E TRABALHOS FUTUROS

Este TG apresentou um método de aquisição de casos de autenticação de data de pinturas. Além disto, para que fosse possível a aquisição destes casos, um padrão de representação dos casos neste problema de aplicação foi desenvolvido. Dentre outros aspectos, a solução proposta permite que os casos de autenticação sejam armazenados em duas bases de casos distintas do sistema e, posteriormente, sejam utilizados para auxiliar um especialista em autenticação de pinturas na investigação de uma solução para novos problemas de autenticação.

As datas de uma pintura, as quais representam as soluções dos sistemas de CBR testados neste TG, eram expressas de diferentes formas. Logo, um padrão para a representação destas datas no sistema de CBR foi proposto e implementado. Com intuito de auxiliar o especialista durante o processo de autenticação de datas de pinturas, o padrão desenvolvido busca seguir os termos encontrados com maior frequência nos relatórios de autenticação disponíveis para análise.

Com o objetivo principal de auxiliar o processo de autenticação de pinturas, a possibilidade da utilização conjunta de pigmentos e de composições elementares foi investigada neste TG. Em particular, o método de combinação proposto visou combinar as similaridades obtidas em cada um dos processos de recuperação, atribuindo uma maior importância para os resultados baseados em pigmentos gerados pelo sistema AuthenticPainting. Contudo, os testes realizados com este sistema permitem observar que este método de combinação de sistemas de CBR foi aplicável em apenas 1/3 dos casos de teste. Isto se deve ao fato de que para gerar um resultado de combinação com similaridade superior a 50%, os resultados de recuperação de casos baseados em pigmentos e baseados em composições elementares ser fundamentados em casos em comum entre os quatro casos mais similares retornados como resposta para consultas.

Por fim, espera-se que o sistema AuthenticPainting venha a suportar a aquisição de novos casos de autenticação de pinturas junto a especialistas neste domínio de aplicação, assim expandindo a base de casos atualmente disponível para o desenvolvimento de projetos de pesquisa futuros. Dentre estes projetos, podemos citar a investigação de outros métodos de combinação de resultados de recuperação de casos em sistemas baseados em múltiplas bases de casos (ou múltiplos sistemas de CBR), juntamente com o uso de outras características relevantes (além de pigmentos e composições elementares) para o processo de autenticação de pinturas. Tal investigação futura pode vir a gerar uma contribuição ainda maior tanto para a

área de autenticação de pinturas (tomada como uma área de pesquisa por si só), quanto para a área de Raciocínio Baseado em Casos.

REFERÊNCIAS

AAMONDT, A.; PLAZA, E. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. **AI Communications**, v. 7, n. 1, p. 39–59, 1994.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. I ed. Madison: Springer, 2006. v. 4p. 738

CHOU, J. Web-based CBR system applied to early cost budgeting for pavement maintenance project. **Expert Systems with Applications**, v. 36, p. 2947–2960, 2009.

DE MANTARAS, R. et al. **Machine Learning and Its Applications**. III ed. Berlin: Springer, 2001. v. 2049p. 325

EASTAUGH, N. Scientific Dating Paintings Eastaugh. **infocus Magazine**, p. 30–39, 2006.

EASTAUGH, N. **Pigment Compendium: A Dictionary and Optical Microscopy of Historic Pigments**. I ed. Butterworth-Heinemann, 2008. p. 960

GOLDSTEIN, J. et al. **Scanning Electron Microscopy and X-ray Microanalysis**. 3rd. ed. Massachusetts: Springer, 2003. v. 1p. 689

KHOSHGOFTAAR, T. M.; SELIYA, N.; SUNDARESH, N. An empirical study of predicting software faults with case-based reasoning. **Software Quality Journal**, v. 14, n. 2, p. 85–111, jun. 2006.

KOLODNER, J. L. **Case-Based Reasoning**. San Mateo, CA: Morgan Kaufmann Publishers, Inc., 1993. v. 5

LEAKE, D.; SOORIAMURTHI, R. **When Two Case Bases Are Better Than One: Exploiting Multiple Case Bases** (I. Watson, D. W. Aha, Eds.) ICCBR. **Anais...**Vancouver: Springer-Verlag, 2001

LEAKE, D.; SOORIAMURTHI, R. **Managing Multiple Case Bases: Dimensions and Issues** 15 Th International Florida Artificial Intelligence Research Society (Flairs) Conference. **Anais...**Pensacola: AAAI Press, 2002

LOPEZ DE MANTARAS, R. et al. Retrieval, reuse, revision and retention in case-based reasoning. **The Knowledge Engineering Review**, v. 20, n. 03, p. 215, 8 maio. 2006.

NAKASUMI, M. **Credit risk management system on e-Commerce** Proceedings of the 5th international conference on Electronic commerce - ICEC'03. **Anais...**New York, USA: ACM Press, 2003

PLAZA, E.; MCGINTY, L. Distributed case-based reasoning. **The Knowledge Engineering Review**, p. 1–4, 2005.

SILVA, L. A. DE L. et al. **A Case for Numerical Taxonomy in Case-Based Reasoning** Proceeding SBIA'08 Proceedings of the 19th Brazilian Symposium on Artificial Intelligence: Advances in Artificial Intelligence. **Anais...**Springer-Verlag, 2008

SILVA, L. A. DE L. **Enhancement of Case-Based Reasoning through Informal Argumentation, Reasoning Templates and Numerical Taxonomy**. London: University College London, 2010.

SILVA, L. A. DE L. et al. **A Case for Folk Arguments in Case-Based Reasoning** Proceeding ICCBR'10 Proceedings of the 18th international conference on Case-Based Reasoning Research and Development. **Anais...**London: Springer-Verlag, 2010

WATSON, I. **Applying Knowledge Management Watson**. I ed. San Francisco, CA: Morgan Kaufmann, 2003.

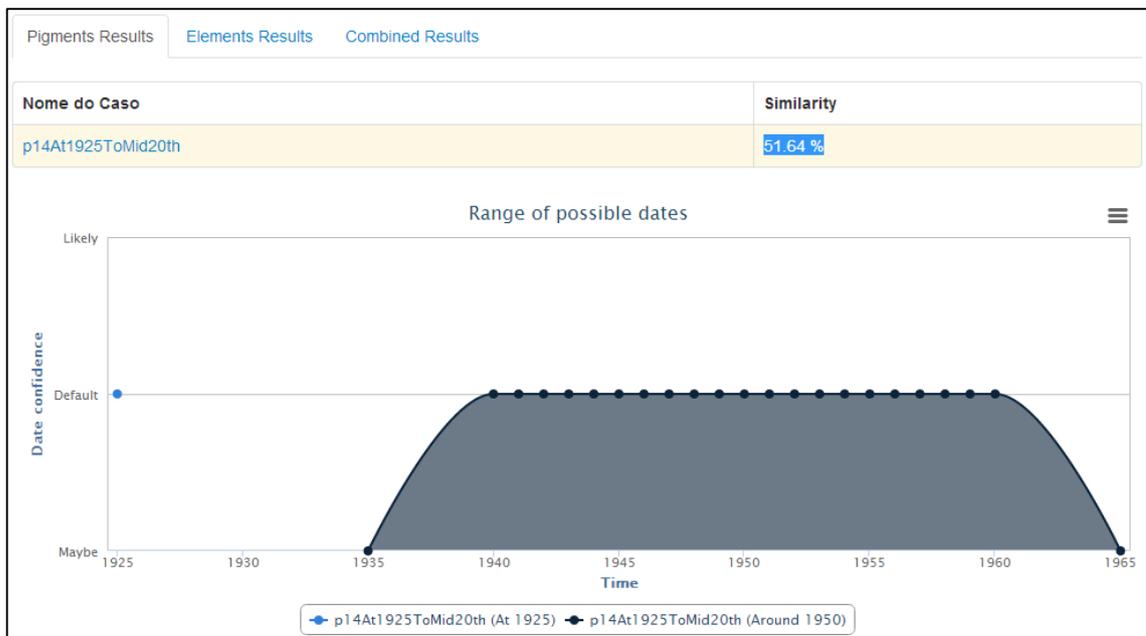
YU, K. A. I.; JI, L.; ZHANG, X. Kernel Nearest-Neighbor Algorithm. **Neural Processing Letters**, v. 15, p. 147–156, 2002.

ANEXOS

Anexo A – Resultados do teste do sistema

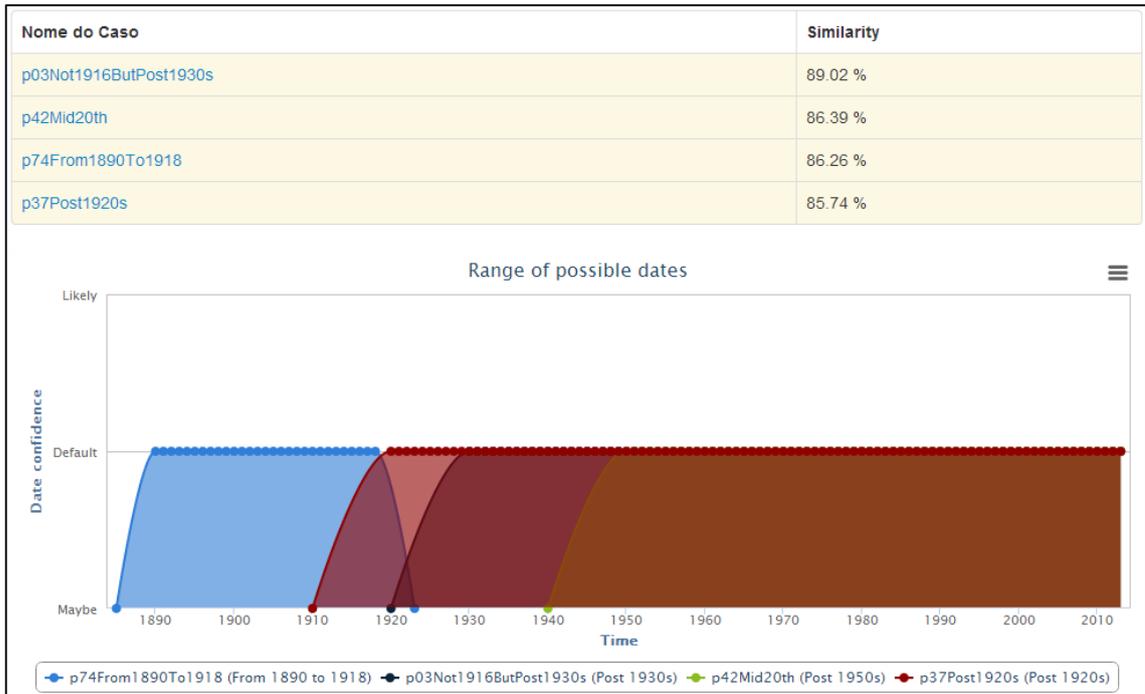
1. Resultados dos testes da pintura p02Not1915ButPost1920to1930

Para a pintura de teste p02, temos o período entre 1920 e 1930 como a data autenticada. Além disto, sabe-se que a pintura não é do ano de 1915. Para esta pintura, não foi possível gerar resultados com similaridade superior à 50% através do método de combinação proposto. As figuras abaixo apresentam os resultados juntamente com uma análise e justificativa.



Resultados da consulta por pigmentos para a pintura teste p02

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
A pintura apresentada encontra-se na faixa de datas da pintura de teste p02.				

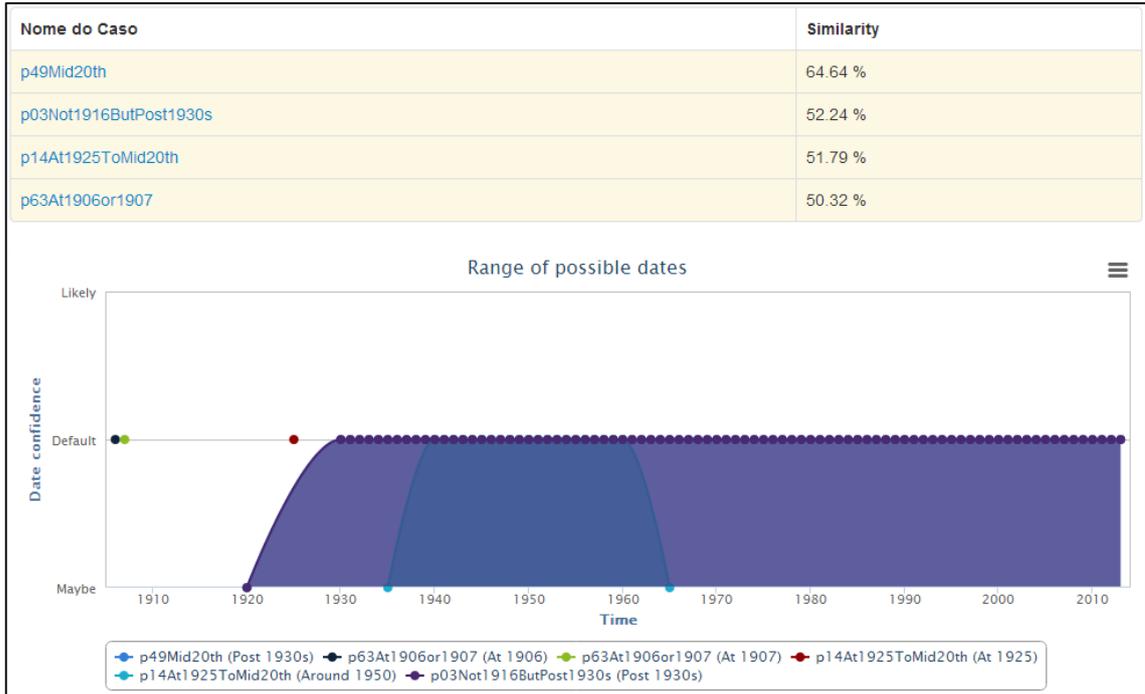


Resultados da consulta utilizando composição elementar para a pintura p02.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p02.				

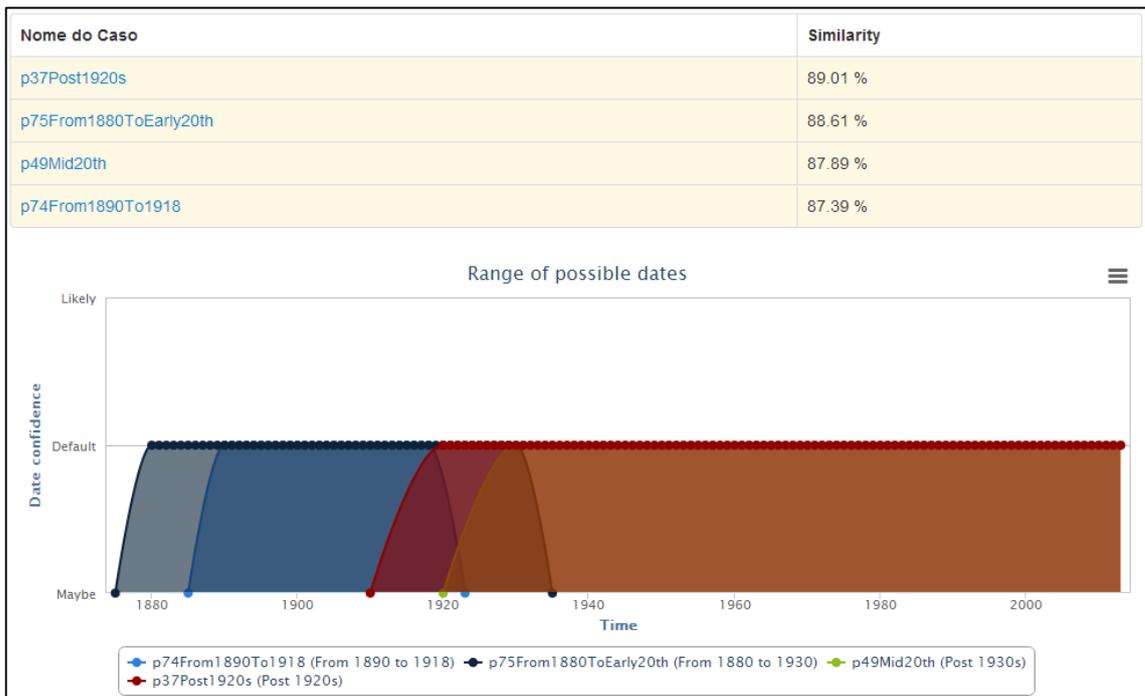
2. Resultados dos testes da pintura p07NotPost1916ButMid1930s

Para a pintura de teste p07, temos como a metade da década de 1930 como a data autenticada. As figuras abaixo apresentam os resultados juntamente com uma análise e justificativa.



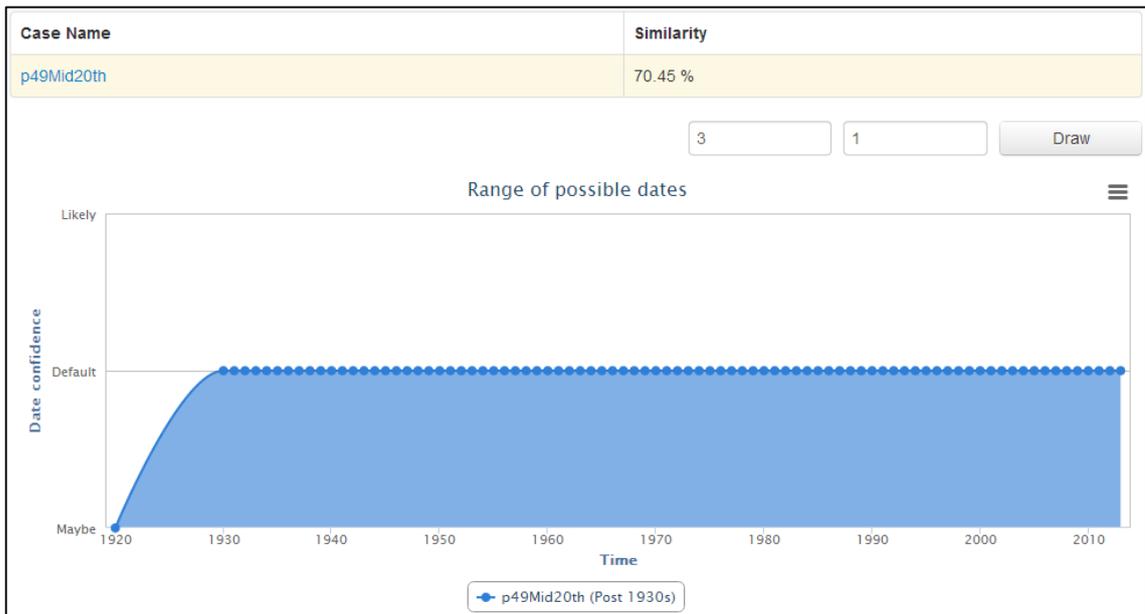
Resultados da consulta utilizando pigmentos para a pintura p07.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas possuem as datas bastante próximas a data real da pintura teste p07.				



Resultados da consulta utilizando composição elementar para a pintura p07.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas possuem as datas bastante próximas a data real da pintura teste p07.				

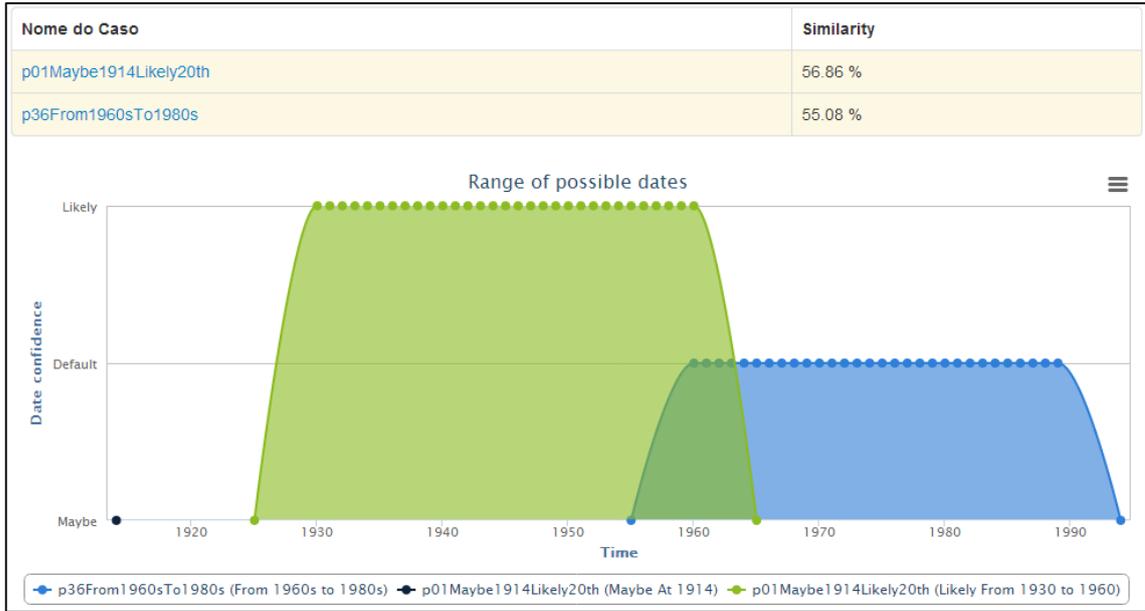


Resultados da combinação utilizando peso 3 para os pigmentos da pintura p07.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas possuem as datas bastante próximas a data real da pintura teste p07.				

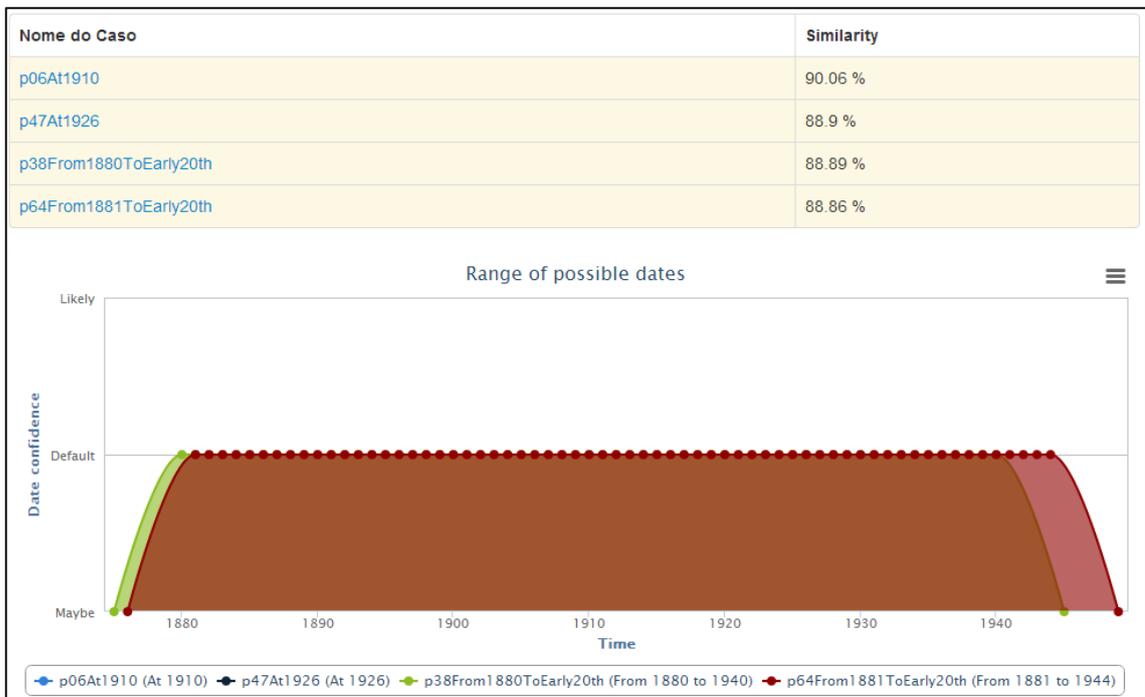
3. Resultados dos testes da pintura p08Around1918

Para a pintura de teste p08, temos as data em torno do ano de 1918 como datas autenticadas pelo especialista. Para este teste, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.



Resultados da consulta utilizando pigmentos para a pintura p08.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
Apenas a primeira pintura recuperada possui uma data utilizável para a solução da pintura teste p08.				



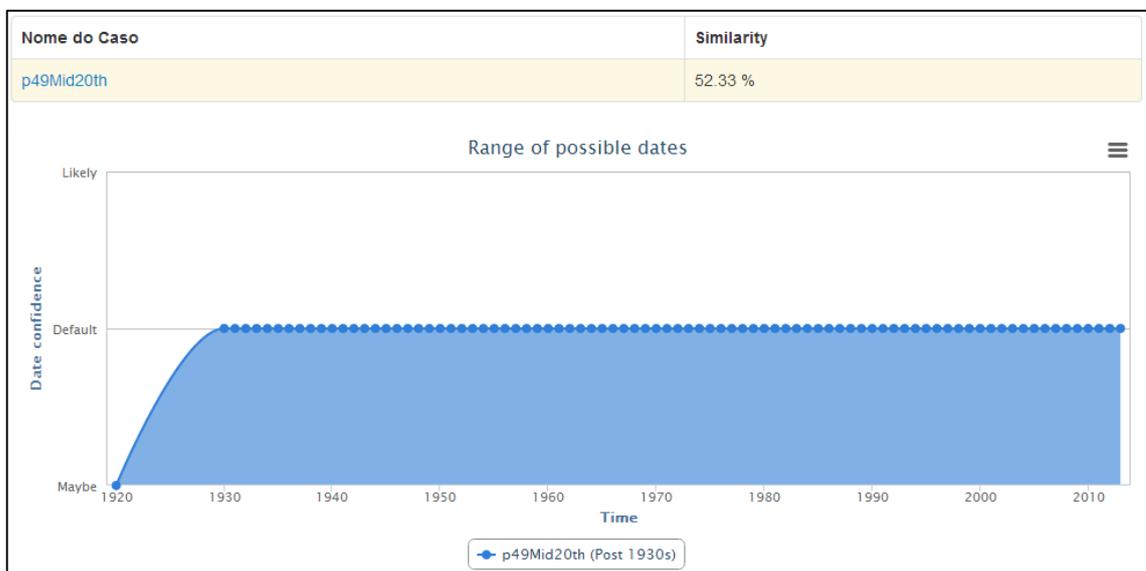
Resultado da consulta utilizando composições elementares para a pintura p08.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável

	X			
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p08, com alta similaridade.				

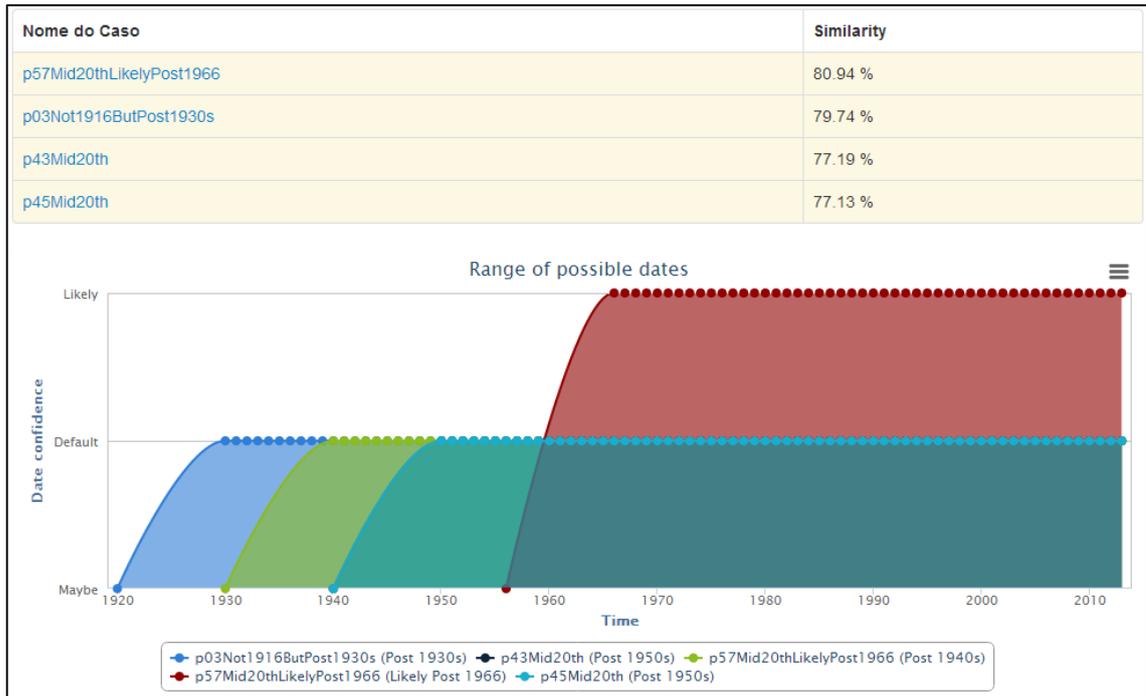
4. Resultados dos testes da pintura p16NotEarly1799But2ndHalf20th

Para a pintura de teste p16, temos a segunda metade do século 20 como a data autenticada pelo especialista. Para esta pintura, o método de combinação proposto não recuperou casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para a pintura p16.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
A pintura recuperada apresenta data próxima a pintura de teste p16, porém com baixa similaridade.				



Resultado da consulta utilizando composição elementar para a pintura p16.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas apresentam datas bastante próximas as datas reais da pintura teste p16.				

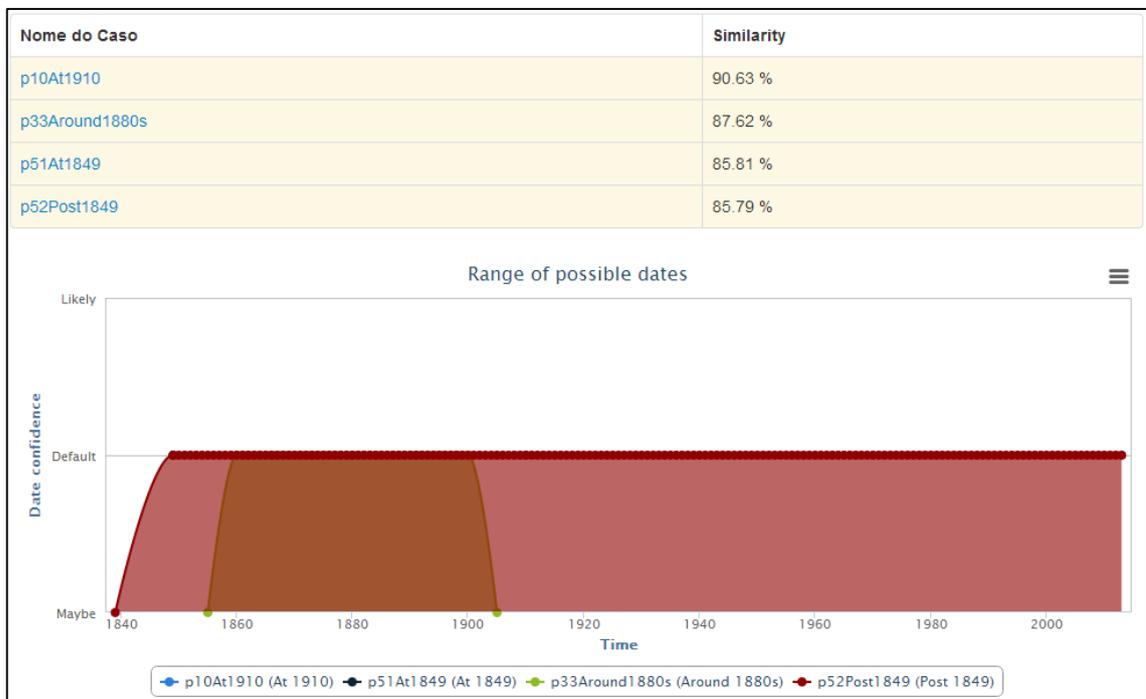
5. Resultados dos testes da pintura p23NotPost1866to1906ButLate20th

Para a pintura teste p23, temos o final do século XX como as datas autenticadas pelo especialista. Para esta pintura, o método de combinação proposto foi capaz de recuperar casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para a pintura p23.

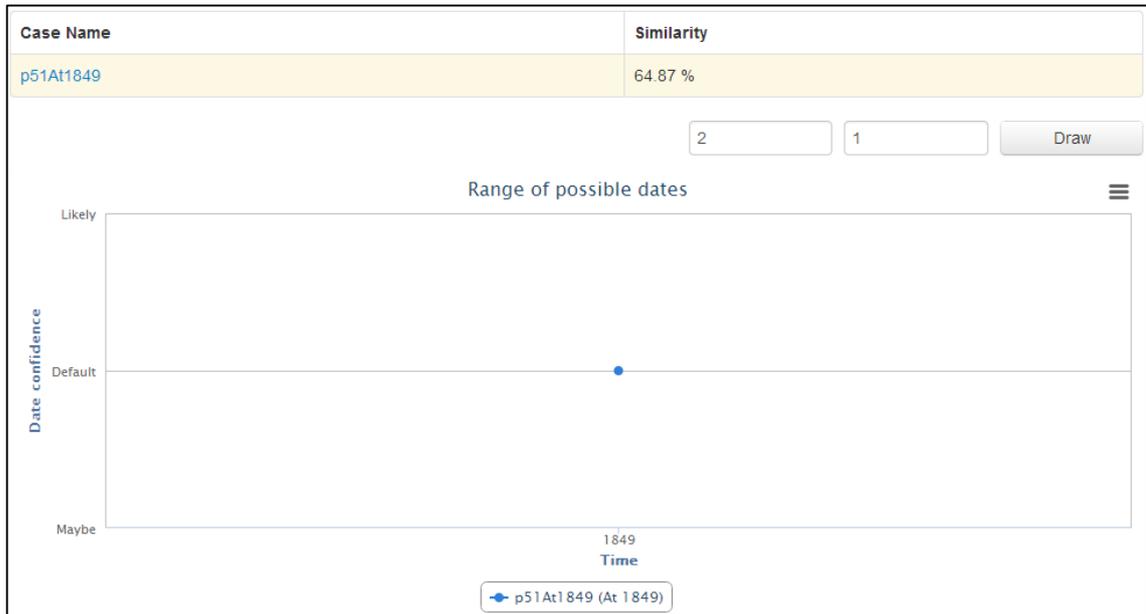
Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
A pintura recuperada apresenta data próxima a pintura de teste p16, porém com baixa similaridade.				



Resultado da consulta utilizando composição elementar para a pintura p23.

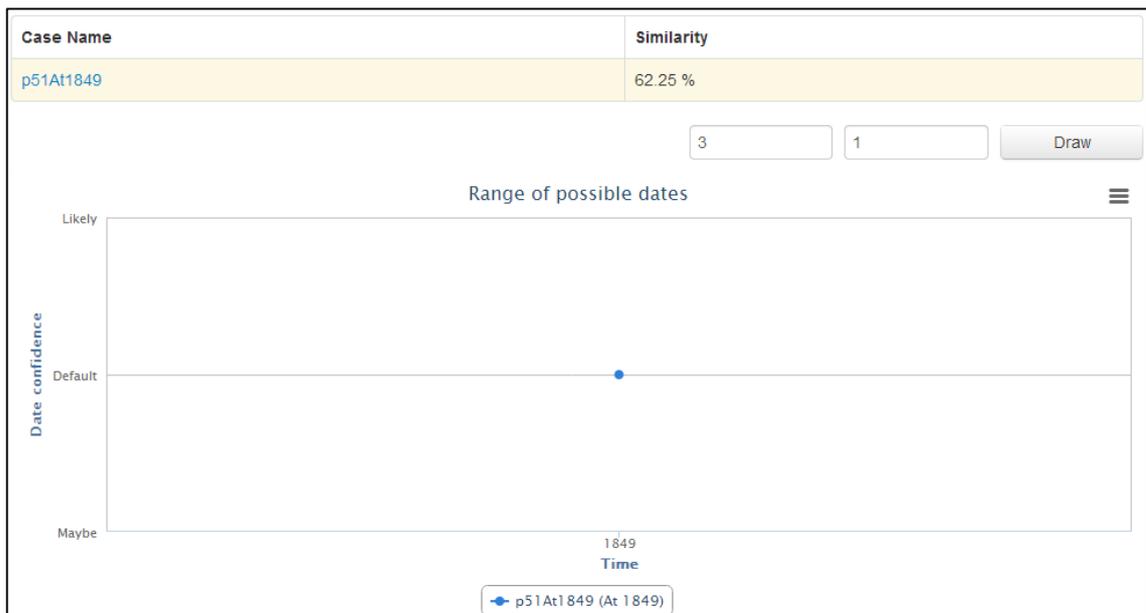
Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		

As duas primeiras pinturas recuperadas apresentam datas próximas das datas reais da pintura teste p23.



Resultado da combinação utilizando peso 2 para os pigmentos.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
A pintura recuperada apresenta datas próximas as datas reais da pintura teste p23, porém a pintura p23 possui uma gama maior de possibilidades de datas.				

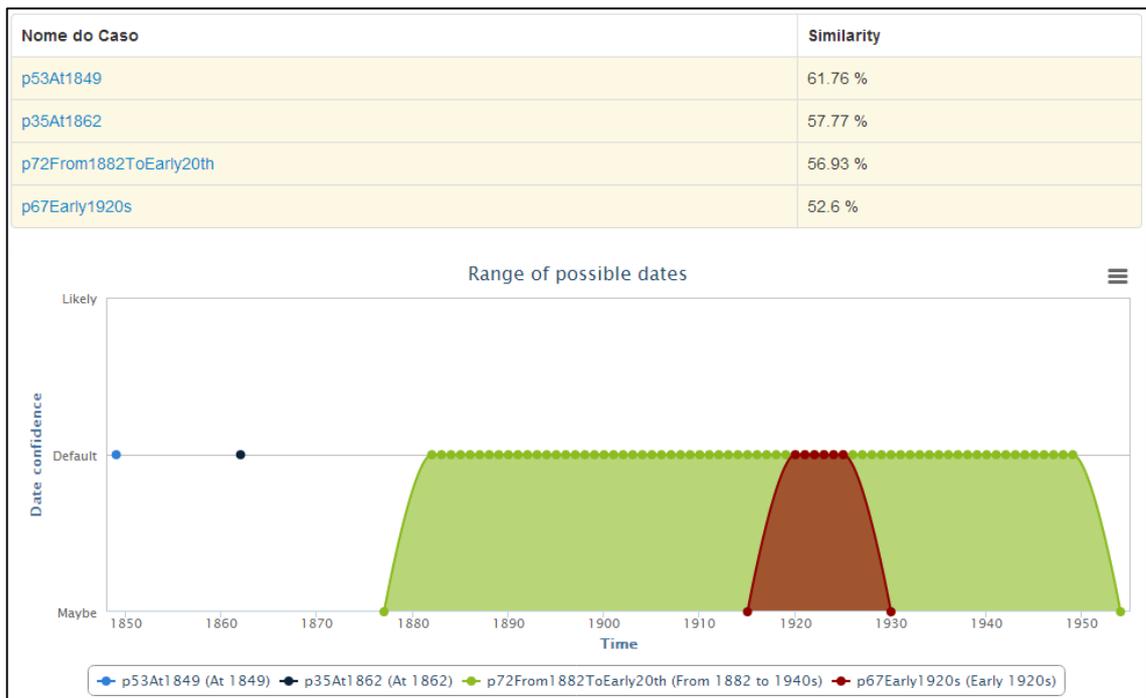


Resultado da combinação utilizando peso 3 para os pigmentos.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
A pintura recuperada apresenta datas próximas as datas reais da pintura teste p23, porém a pintura p23 possui uma gama maior de possibilidades de datas.				

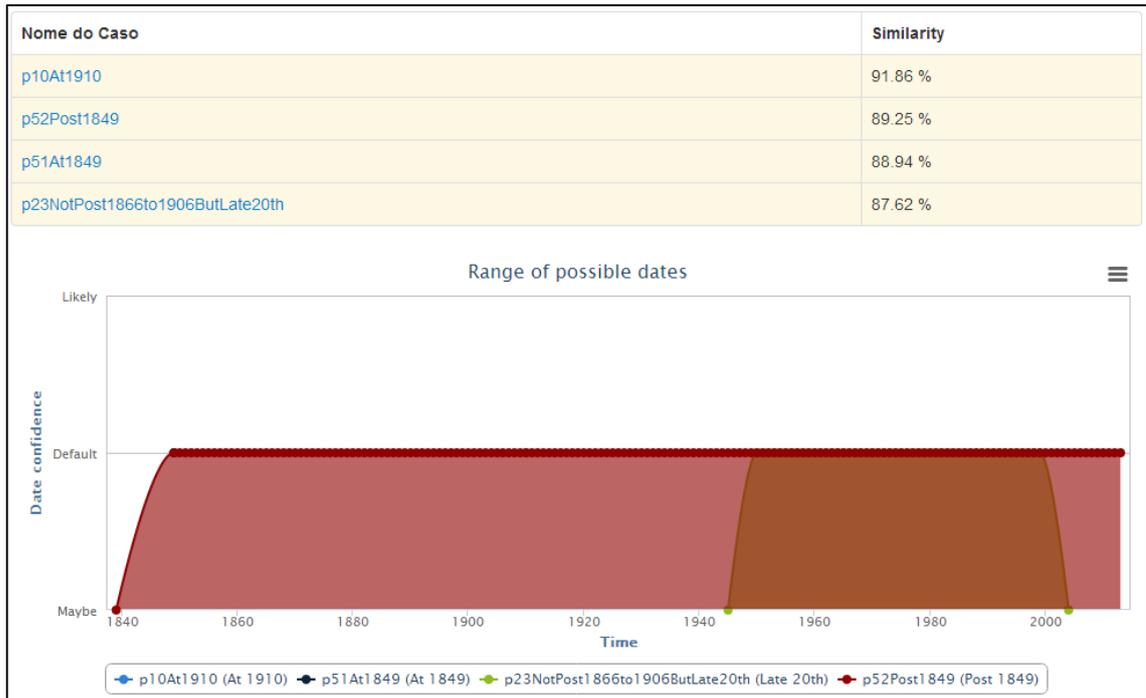
6. Resultados dos testes da pintura p33 Around 1880s

Para a pintura teste p33, temos o período de datas próximas a década de 1880 como as datas autenticadas pelo especialista. Para esta pintura, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para a pintura p33.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
As pinturas recuperadas possuem as datas concentradas próximas a data real da pintura teste p33.				

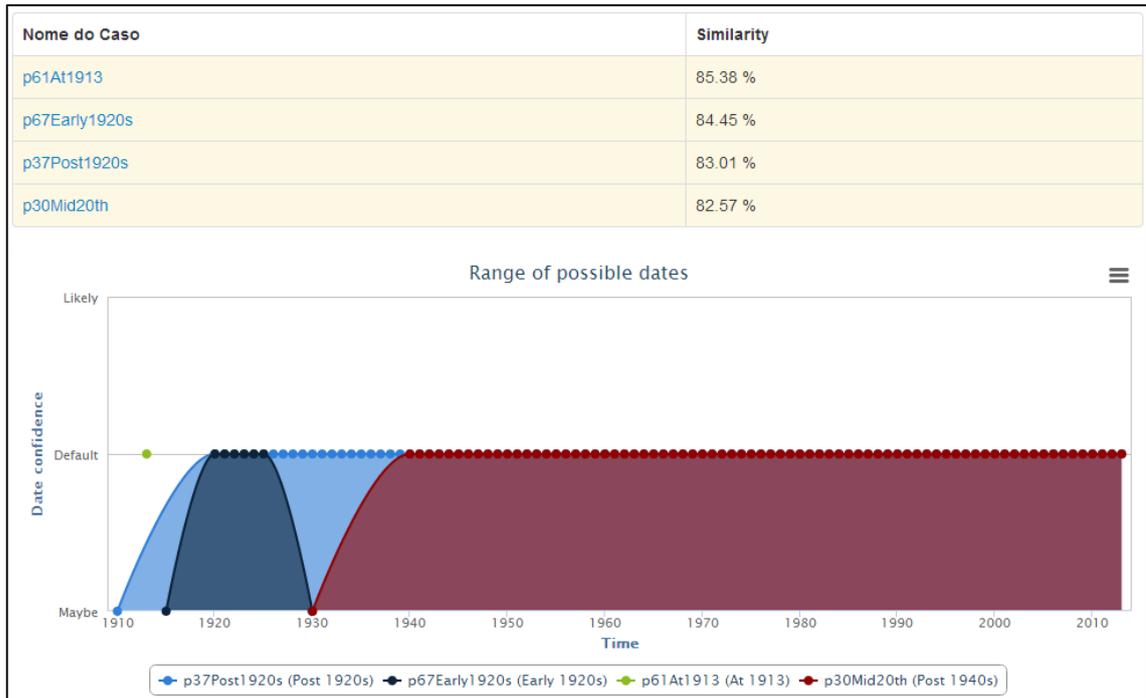


Resultado da consulta utilizando composição elementar para a pintura p33.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
Apenas algumas pinturas apresentam datas próximas das datas reais da pintura teste p33.				

7. Resultados dos testes da pintura p50Mid20th

Para a pintura p50, temos a metade do século XX como as datas autenticadas pelo especialista. Para esta pintura, não foram recuperados casos quando os pigmentos foram utilizados como consulta. Por consequência, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.

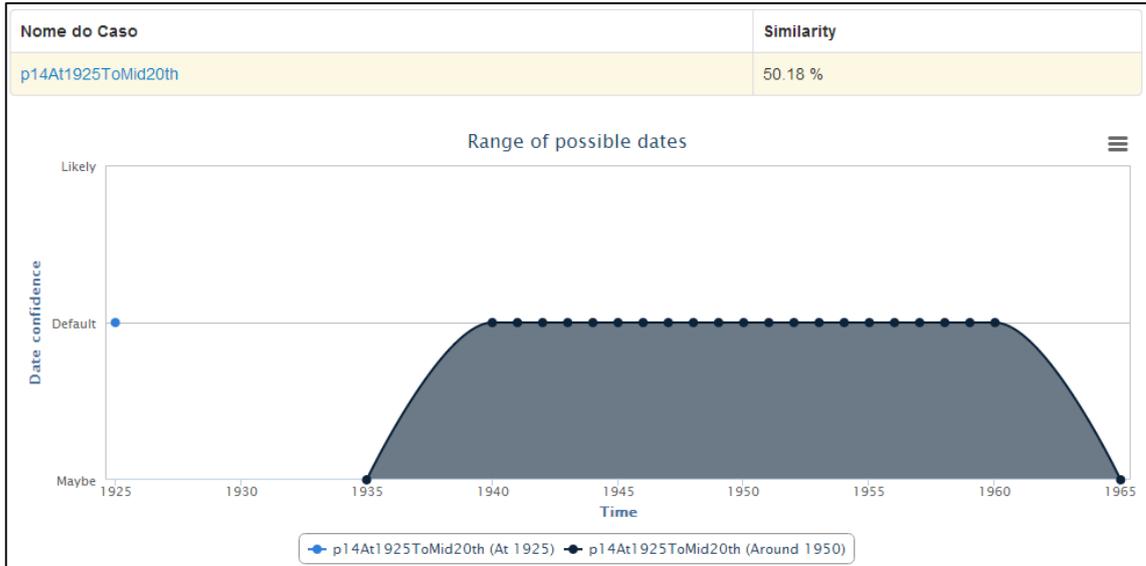


Resultado da consulta utilizando composição elementar para a pintura p50.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p50.				

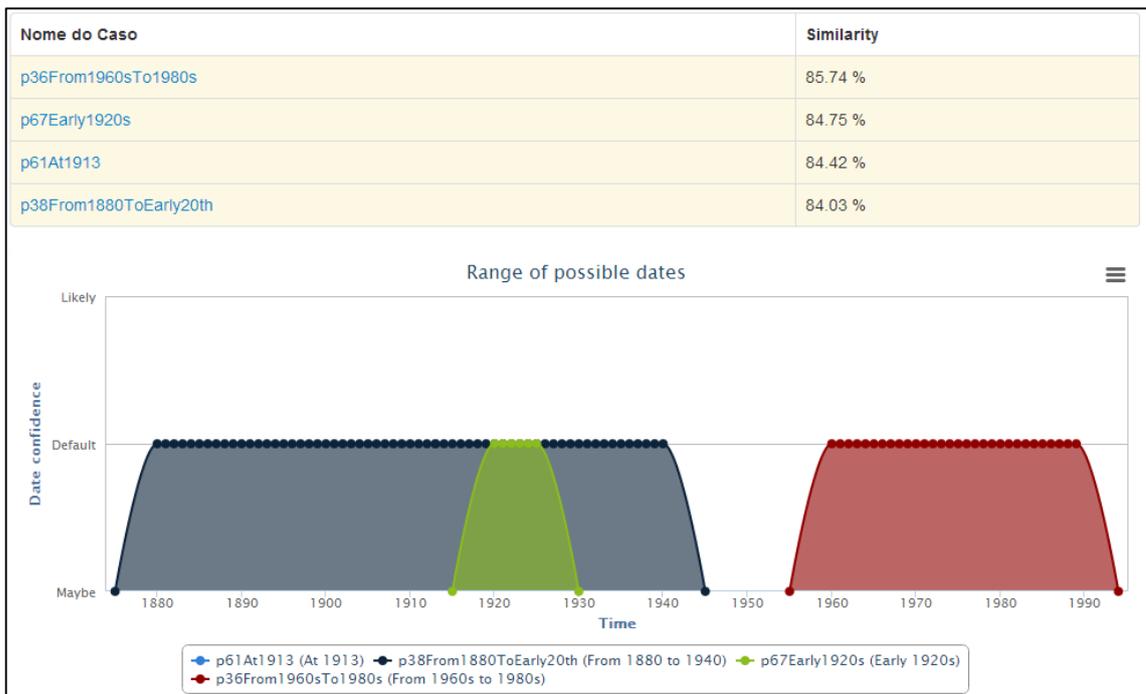
8. Resultados dos testes para a pintura p58Not1914ButMid20th

Para a pintura p58, temos a metade do século XX como as datas autenticadas pelo especialista. Para esta pintura, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para pintura p58.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			
A pintura recuperada apresenta datas próximas as datas reais da pintura teste p58.				



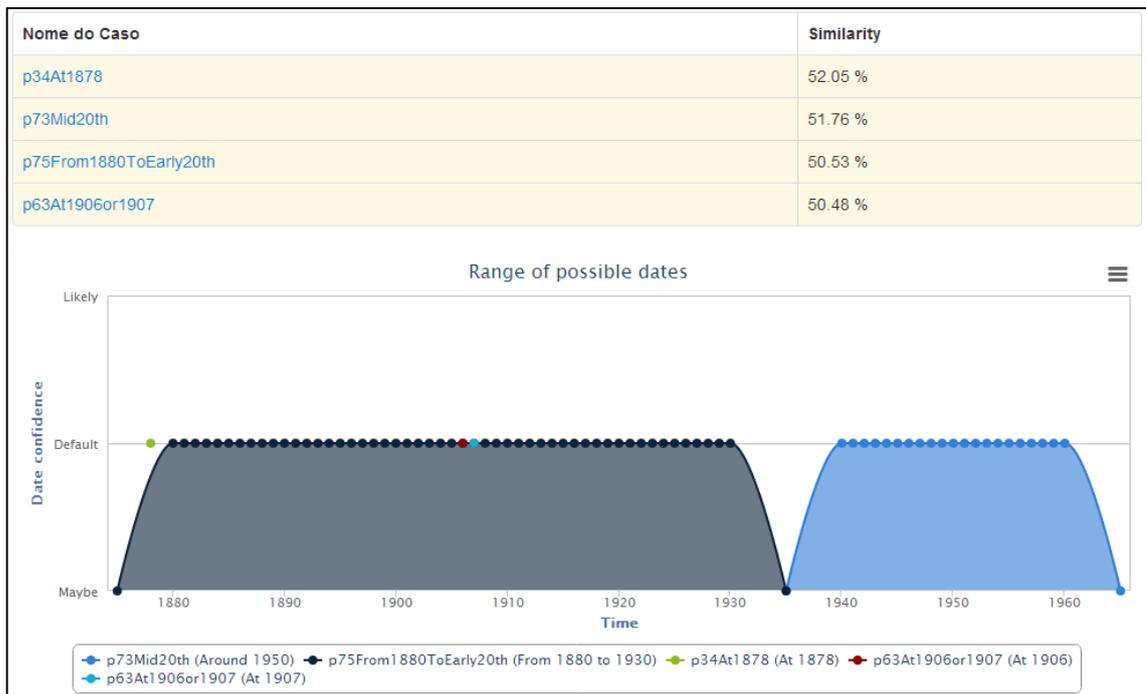
Resultado da consulta utilizando composição elementar para a pintura 58.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		

As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p58.

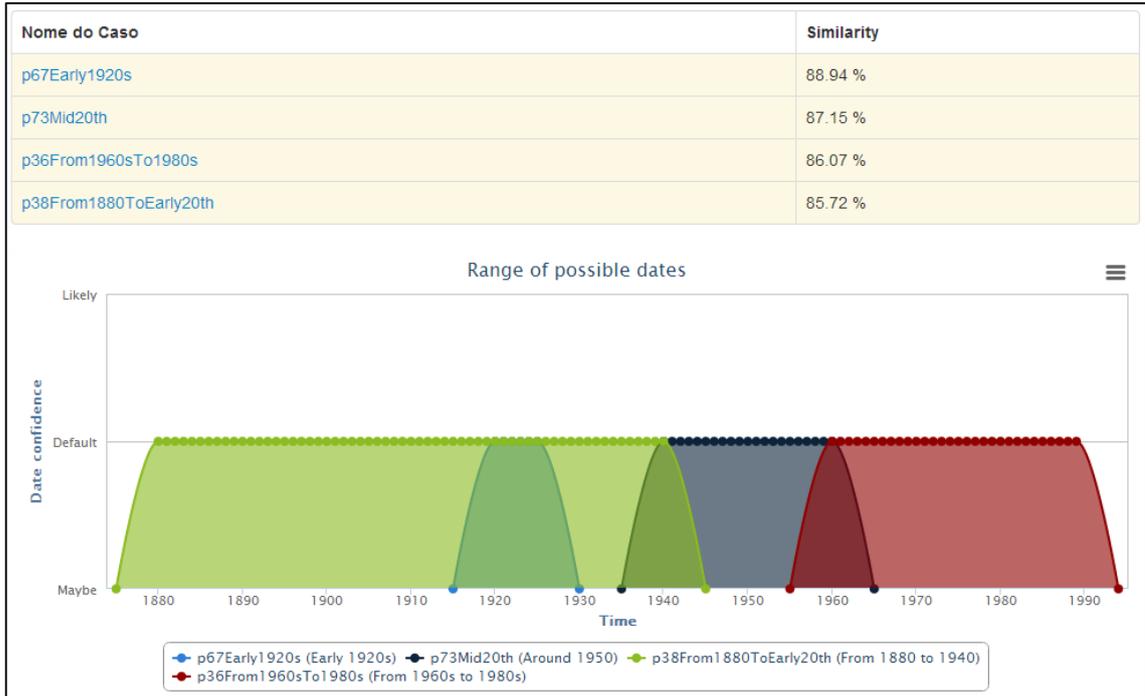
9. Resultados dos testes da pintura p61At1913

Para a pintura p61, temos o ano de 1913 como a data autenticada pelo especialista. Para esta pintura, o método de combinação proposto foi capaz de recuperar casos similares com similaridade superior a 50%.



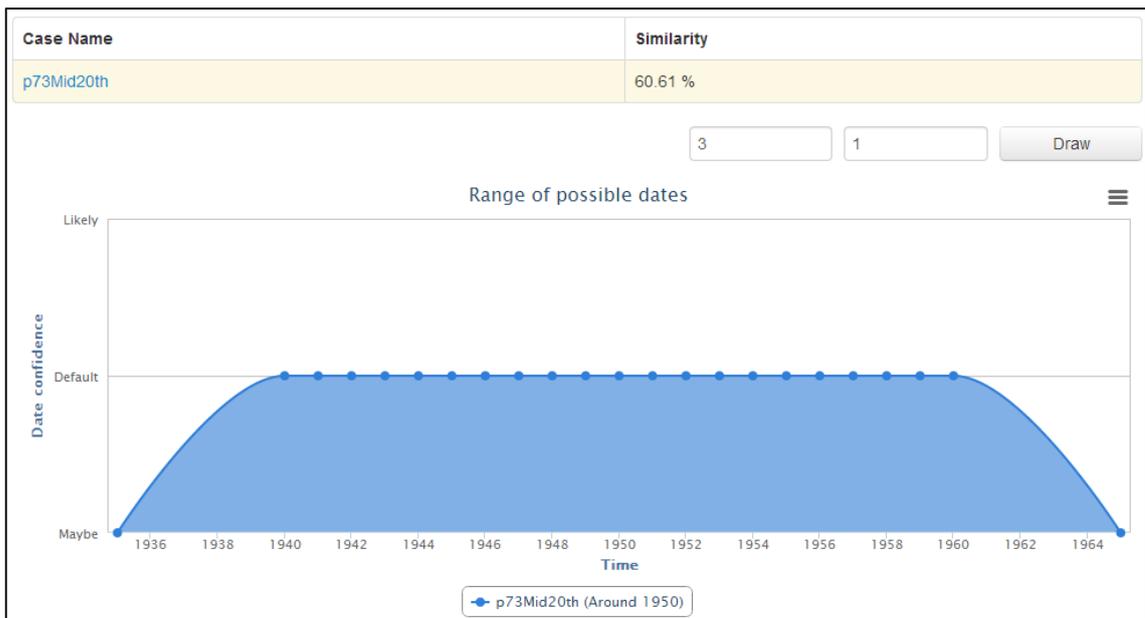
Resultado da consulta utilizando pigmentos para a pintura p61.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
Algumas das pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p61.				



Resultado da consulta utilizando composição elementar para a pintura p61.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
		X		
As duas primeiras pinturas recuperadas apresentam datas próximas das datas reais da pintura teste p61.				



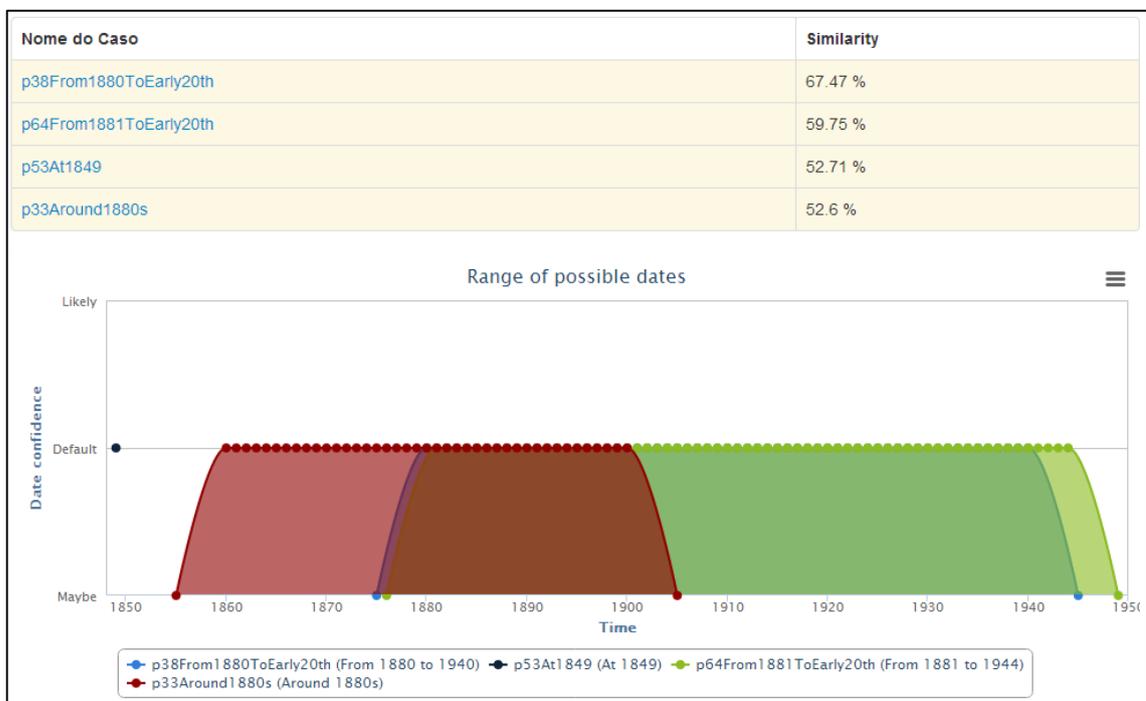
Resultado da combinação utilizando peso 3 para os pigmentos.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
---------------	------	-------------	----------	---------------

X				
A pintura recuperada apresenta datas bastante próximas as datas reais da pintura teste p61 para ambos os pesos de pigmentos.				

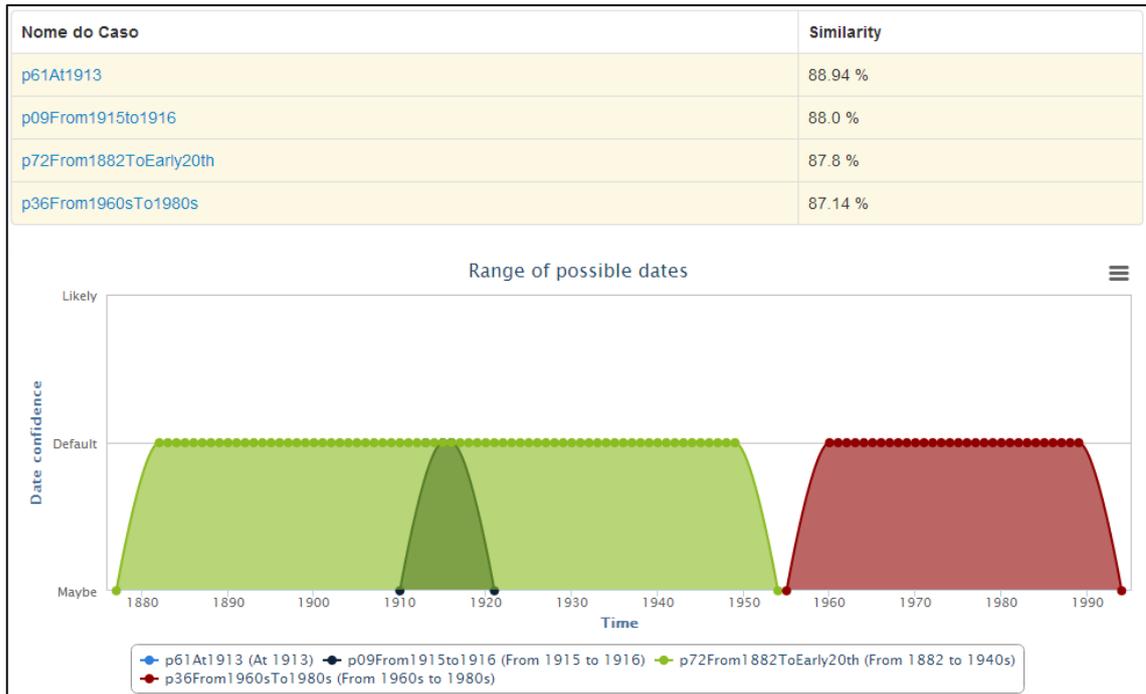
10. Resultados dos testes da pintura p67Early1920s

Para a pintura p67, temos o começo da década de 1920 como as datas autenticadas pelo especialista. Para esta pintura, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para a pintura p67.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p67, todas com alta similaridade				

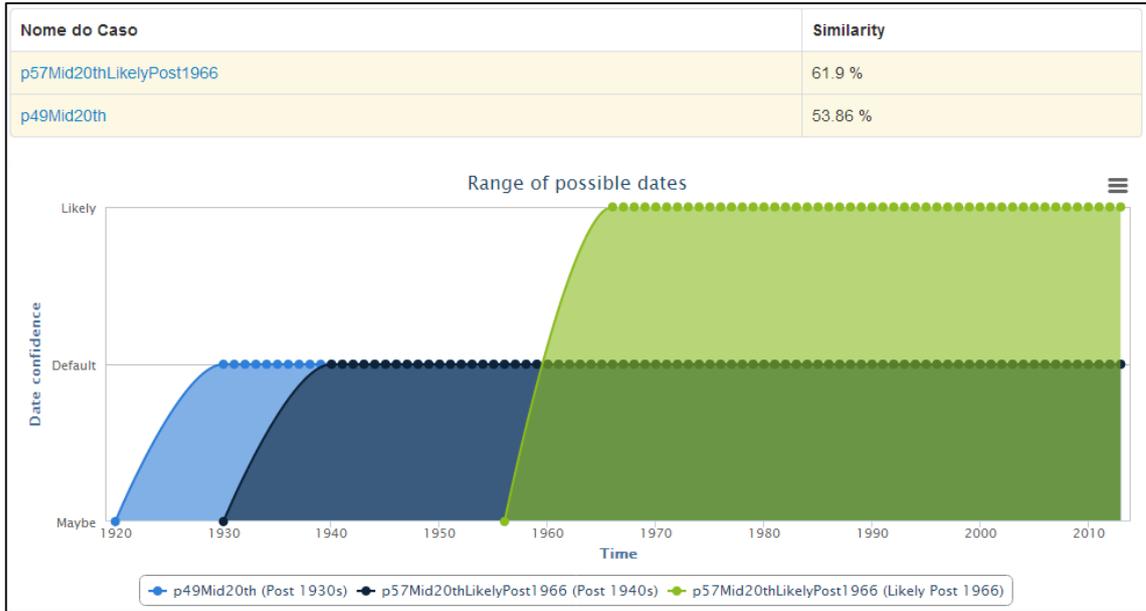


Resultado da consulta utilizando elementos para a pintura p67.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
X				
As pinturas recuperadas apresentam datas bastante próximas as datas reais da pintura teste p67.				

11. Resultados dos testes da pintura p76Not1913ButMid20th

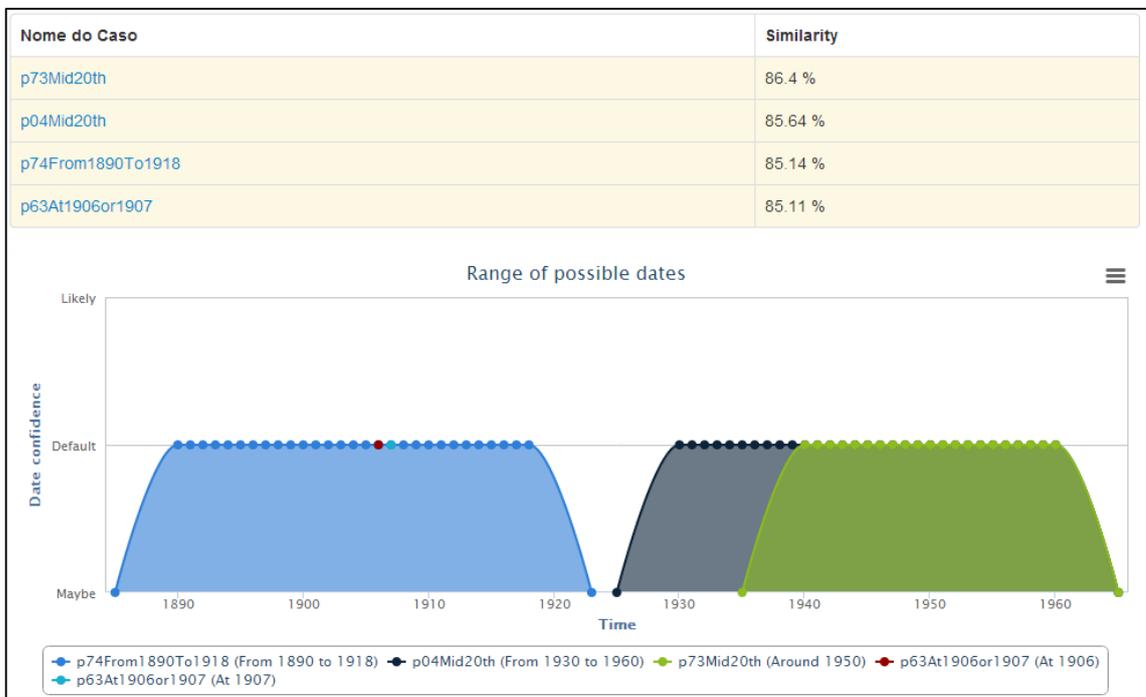
Para a pintura p76, temos a metade do século XX como as datas autenticadas pelo especialista. Para esta pintura, o método de combinação proposto não foi capaz de recuperar casos com similaridade superior à 50%.



Resultado da consulta utilizando pigmentos para a pintura p76.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável
	X			

As pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p76, respeitando as datas que não são aceitas para a pintura teste.



Resultado da consulta utilizando composição elementar para a pintura p76.

Bastante útil	Útil	Talvez útil	Não útil	Não aplicável

		X		
As duas primeiras pinturas recuperadas apresentam datas próximas as datas reais da pintura teste p76.				

APÊNDICES

Apêndice A – Um relatório de autenticação de pinturas

Introduction

This report presents the results of a scientific examination of the painting *In the Dukhan, Imaginary Journey to Turkey*, thought to be by Mikhail Larionov and to date from 1910. The painting measures 50cm (h) by 70cm (w) and has been executed on canvas.

The purpose of this report is to provide an analysis of pigments present in the painting with the aim of establishing whether there is reason to doubt the dating and attribution of the work. To this end a series of samples was removed for analysis, following standard methods for identification of pigments. Additionally, some information on the materials and techniques employed by Larionov that is available was used to compare with the results of the analysis.

From the analyses and comparative studies, no pigments anachronistic for a date of 1910 were found. Moreover two pigments specifically identified as present in the painting are likely to be positively indicative of a work from the supposed date of the painting. It was also found that the types of pigments used are very similar to those that Larionov is known to have employed in the period.

Analysis

Prior to sampling, the painting was examined in detail visually, with the aid of low-power binocular microscopy, and with UV fluorescence. Little evidence of restoration was apparent, and the painting is otherwise generally in a condition not inappropriate to its supposed age. (However, such subjective judgments of condition as related to chronological age are not necessarily sufficiently reliable or discriminatory for firm assertions regarding authenticity.) The painting appears to be still mounted on its original 4-piece adjustable wooden stretcher.

The range of pigments present was identified from a series of twelve samples taken from a variety of locations on the painting. All samples were taken from layers that appeared to be original (that is, were clearly contiguous with those below them). Locations were further selected so as to represent as wide a range of the colours (and therefore probably pigments also) as possible. Thus the pigments identified and discussed below represent, as far as can be determined, the extent of the original palette used by the artist; however, the presence of further pigments within the painting cannot be wholly ruled out.

The samples were subjected to analysis by polarised light microscopy ('PLM'), scanning electron microscopy-energy dispersive X-ray spectrometry ('SEM-EDX') and UV-visible-near infrared micro-spectrophotometry. All are standard methods within the field,

capable of allowing the kinds of differentiation required for this type of work. Comparison was also made between samples from the painting and examples of similar pigments from a large collection of reference standards previously analysed by multiple means. Certain differentiations cannot be made from this range of techniques, although for present purposes the level of discrimination would appear to be sufficient. For example, different types of lead carbonate properly require other methods such as Raman microscopy, but the implications of finding lead carbonate or lead carbonate hydroxide are minimal. However, it is suggested that if further confirmation is required, then such additional analysis could be conducted.

The following pigments were identified: synthetic ultramarine, iron(III) hexacyanoferrate(II) ('Prussian blue'), manganese phosphate ('manganese violet'), 'dry' process mercury(II) sulfide ('vermilion'), copper di-potassium hexacyanoferrate(II) ('Hatchett's brown' and other terms), lead chromate ('chrome yellow'), a lead-based white, probably as lead carbonate hydroxide ('lead white'), barium sulfate, calcite, an unidentified red 'lake' pigment and an earth pigment. There was also what is probably a compound pigment present, that sometimes (misleadingly) known as 'chrome green', a co-precipitate of yellow lead chromate and blue iron hexacyanoferrate(II). The barium sulfate is likely to be an 'extender' pigment (that is, one used in the manufacture of satisfactory paints as opposed to something specifically selected by the artist).

All pigments were generally identified through a combination of the techniques applied. However, certain key features were specifically determined through one or other method. For example, the particle morphology of mercury(II) sulfide is clearly determined through the PLM, this indicating the 'dry' process involving sublimation rather than derivation by an aqueous route (the 'wet' process). Manganese phosphate is a member of a group of violet pigments that can be quite well differentiated by PLM, with elemental analysis essentially providing support for the result; it appears to be the manganese(III) phosphate rather than the alternate ammonium manganese phosphate, though additional analysis could be applied to wholly confirm this. Ultramarine was differentiated from pigments of similar optical appearance by elemental analysis (presence of key elements such as aluminium, silicon and sulfur, with an absence of elements such as cobalt that might indicate cobalt aluminium oxide) and measurement of optical spectra (this showing high transmission of red/near infrared and an absence of the characteristic cobalt absorption bands). Iron(III) hexacyanoferrate(II) was confirmed through optical appearance, presence of weak peaks for iron in relevant samples and the spectral transmission (notably low transmission in the red/near infrared). Lead chromate was evident by both PLM and SEM-EDX; it appears to be lead chromate(VI), though this should properly be confirmed

by other techniques such as Raman microscopy since there is some chronological differentiation among the related group of compounds. The 'chrome green' (lead chromate/iron hexacyanoferrate(II) combination) was identified on the basis of the intimate association of the two phases, suggesting a co-precipitated mixture. The lead carbonate-based white was closely similar to reference material by PLM, but the elemental analysis essentially confirmed the identity and apparent absence of other phases. Barium sulfate was identified wholly on the basis of SEM-EDX as there is relatively little of it present and particle characteristics were not clearly visible for robust identification by PLM. The copper di-potassium hexacyanoferrate(II) pigment was closely comparable by PLM to reference samples of this compound prepared by this author and colleagues; the presence of copper, iron and potassium by SEM-EDX was confirmatory.

Discussion

No pigments inappropriate for a date of 1910 were found. For example, numerous false paintings in the style of the Russian *Avant Garde* movement contain titanium dioxide white pigments, for which current evidence suggests there was little general use before the 1950s. This pigment was not detected, nor others like it.

Several pigments, though still used today, have largely fallen from use. Of these, the presence of a lead-based white pigment and the 'dry' process mercury(II) sulfide are of some interest. In the former case we find around this time the introduction of so-called 'leaded zinc whites', which are found extensively in Russian *Avant Garde* paintings from the second decade of the twentieth century. In a similar vein there has been a shift in production of mercury(II) sulfide as a pigment during the nineteenth and twentieth centuries. Historically the principal method was the 'dry' form, but it seems that effective methods of preparing this pigment aqueously were developed in the nineteenth century. Although the dry form continued to be made, there was a gradual cross-over to the wet method, such that this became the predominant type in the twentieth century before it fell substantially from use with the availability of viable alternatives. Although this is less strong an indicator of date than others mentioned here, the use of dry-process mercury(II) sulfide in the painting is broadly supportive of an earlier dating.

Detection of manganese phosphate is important to the probable dating of the painting. Several manganese phosphates are listed in the historical pigment literature, notably ammonium manganese phosphate ($(\text{NH}_4)_2\text{Mn}_2(\text{P}_2\text{O}_7)_2$), and manganese(III) phosphate. There are various claims as to the date of origin: the British colourmen Winsor & Newton listed 'permanent violet' and 'permanent mauve' in 1892, stating in their 1896 catalogue that the former was 'Phosphate of Manganese'; a later catalogue of theirs (1930) seemingly claims they introduced

it in 1890. However, Rose states it was first prepared by E. Leykauf in 1868 under the term 'Nürnberg violet'. Whatever the truth, it was not very widely employed, but was none-the-less clearly in use before the presumed date of the painting of this report. There is also a degree of confusion regarding the historical composition of this pigment, with a number of sources indicating the ammonium form was that typically produced; it seems probable though that both forms were commercially available at that time.

However, importantly, a substantial group of other mauve or violet pigments appeared, such as the various cobalt phosphates and arsenates. These must have had some superiority, since they largely supplanted the manganese violet pigments. An interesting example of this occurs in the study by Aitken *et al* on Larionov and Goncharova previously mentioned; all works 1911-14 listed contain manganese violet, while those dated 1930-54 contain cobalt violet.

Identification of the copper hexacyanoferrate(II) is highly remarkable in that there appear to be no other reported occurrences on paintings described in the modern technical literature at the present time. However, there are a number of historical documentary sources that describe it. In terms of the historical use of this pigment, those sources most relevant include the original papers by Hatchett (1802), Salter (1869), Bersch (1901) and Coffignier (1924); of these, the latter two are substantial texts dealing with manufacture of pigments. There are several later mentions, but these appear to be either confined to comments stating that it is little used (Doerner, 1935) or essentially historical reviews based on much earlier literature (*Colour Index*, 1971), rather than manufacturing recipes. On the basis of this information we can consequently state that a) the pigment was probably very little used and b) the availability of recipes for the pigment in standard texts on pigment manufacture of the early twentieth century would suggest on the other hand that its occurrence in a work of 1910 is not inappropriate.

The use of synthetic ultramarine and an iron hexacyanoferrate(II) pigment are consistent with Larionov's practice. According to Aitken *et al*, among the paintings they examined by Larionov and Goncharova, these artists used four blue pigments: ultramarine, cobalt, cerulean (cobalt tin oxide) and Prussian; cobalt blue was found only in the works by Goncharova (that is, it was not used by Larionov among his works). Cobalt tin oxide was not found in the present painting however.

Some pigments noted by Aitken *et al* were not found in this painting, the most worthy of note probably being the difference in green pigments. In the present case a mixed green formed of lead chromate and iron hexacyanoferrate(II) was found; Aitken *et al* found instances

of chromium oxide hydrate ('viridian') and copper acetate arsenite ('emerald green'). This difference is probably not greatly significant.

Conclusions

This analysis has shown that there are no overt reasons on the basis of the pigments in the painting to suppose that this picture is other than it is currently supposed to be. Apart from the absence of anachronistic materials, the presence of a number of pigments consistent with early twentieth century forms and use supports this contention.

Sampling

Sampling followed common methodologies for removal of material from paintings for analysis, with due care to take the minimum amount of sample necessary from relevant locations in a robust manner.

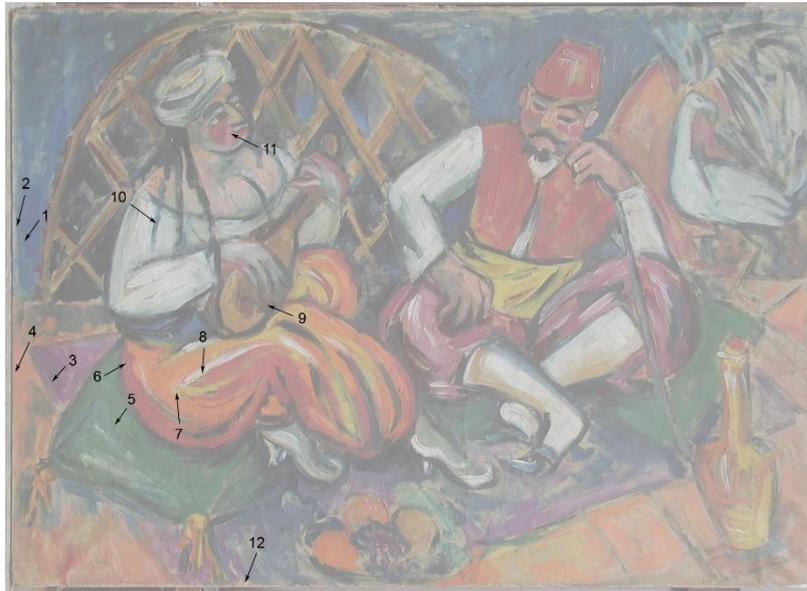
#	<i>Sample Description</i>	<i>Location*</i>
1	Mid-blue	14/292
2	Greenish-blue	9/297
3	Mauve	33/180
...
11	Pink	190/390
12	Off-white ground	203/4

<i>Sample</i>	<i>Colour</i>	<i>Spectrum</i>	<i>Elements**</i>
1	Mid-blue	1	Pb Na Al Si Ba K
2	Greenish-blue	1	Pb Fe
		2	Pb Ca Fe
3	Mauve	1	Pb P Mn K
		2	Pb P Mn Al K
...
11	Pink	1	Pb Al
		2	Pb Al Ca Ba Mg

12	Off-white ground	1	Ca Mg Si
----	------------------	---	----------

*Location is expressed as the distance in millimetres from the left-hand and lower edges (i.e., as x/y co-ordinates).

Elements are listed as **Major, *Minor* and Trace. Carbon and oxygen occurred in all and have been omitted.



Location of samples

