

# Seminar: Multimediale Informationssysteme

SQL/MM

SQL Multimedia and Application Packages

Christian Müller

12.07.2002

## Übersicht

- 1 Einleitung
- 2 Full Text
- 3 Spatial
- 4 Still Image
- 5 Zusammenfassung

## Einleitung

- neue Datentypen für die Speicherung und Verarbeitung multimedialer Daten
- hier behandelt: Text, Bilder, geographische Daten
- kein Audio, Video
- auf Basis der Typerweiterbarkeit von SQL:1999
- UDT (benutzerdefinierte Datentypen) mit eigenen Funktionen und Methoden

## Full Text

- zur Speicherung und Verarbeitung von Texten
- Texte als semantische Einheit (Wörter, Sätze, Absätze)
- Zuordnung von Sprache zu den Texten
- zur Suche in den Texten mehrere Methoden
- verschiedene Arten von Suchmustern (Pattern)

## Methoden zur Suche

```
SELECT name, b.Text.Score(`SQL`)  
FROM Beispiel b  
WHERE b.Text.Contains(`Standard%`);
```

- Contains: prüft auf Existenz des Pattern im Text
- Score: gibt Auskunft darüber wie gut das Pattern zum Text passt
- NumberOfMatches: Anzahl der Vorkommen des Patterns

## einfache Pattern

- Wortsuche:

BeispielText.Contains( ' "Standard" ' )

- Phrasensuche

BeispielText.Contains( ' "internationaler Standard" ' )

- Suche nach mehreren Wörtern / Phrasen

BeispielText.Contains( ' "Standard\_" ' )

BeispielText.Contains( ' "internationaler % Standard" ' )

## Pattern

Pattern zur Suche von Ausdrücken, welche in bestimmten Zusammenhängen mit dem Startausdruck stehen:

- ähnlicher Klang
- ähnliche Buchstabenfolge
- Synonyme
- allgemeinere Ausdrücke
- speziellere Ausdrücke

## Pattern

Suche nach Synonymen von *Liste* im Thesaurus von *Informatik*

```
BeispielText.Contains( `THESAURUS "Informatik"  
EXPAND SYNONYM TERM OF "Liste" `)
```

## Pattern

Suche nach Stammformen

~STEMMED FORM OF "Mäuse" ~ wird zu  
~( "Maus" , "Mäuse" )~

~STEMMED FORM OF „Internationale Standards“ ~ wird zu  
~("International Standard" , "Internationaler Standard" ,  
"Internationales Standard" , ... )~

## Pattern

proximity expansion pattern

```
`(("Standard", "International")  
NEAR "Sprache" WITHIN 0 SENTENCES IN ORDER)`
```

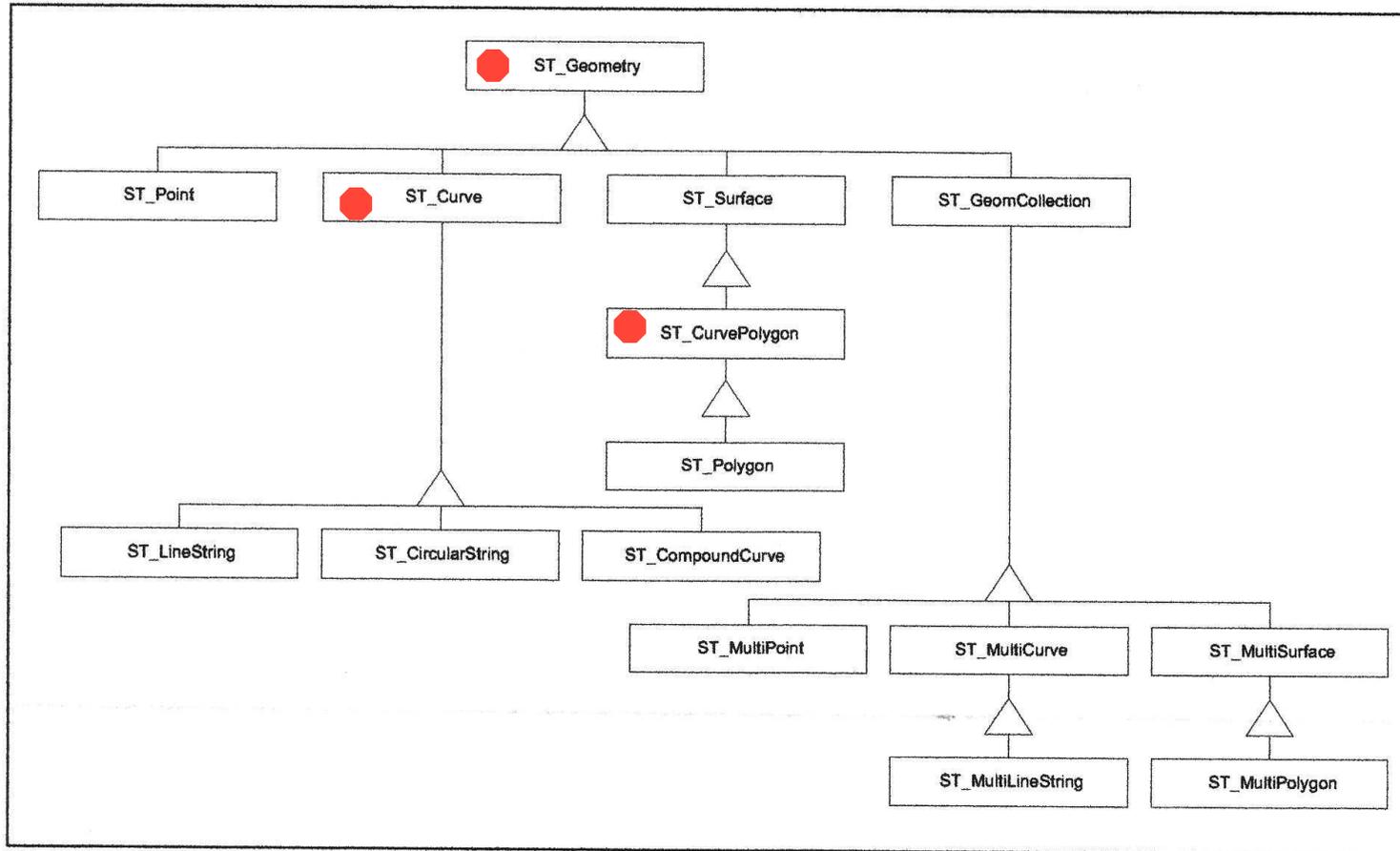
context condition pattern

```
`("Standard", "International")  
IN SAME SENTENCE AS "Spache"`
```

## Spatial

- Datentypen zur Speicherung von geographischen Daten
- 0-, 1- oder 2-dimensionale Daten
- Methoden für die Berechnung von Lagebeziehungen
- Einsatzmöglichkeiten bei Stadtplanung, Landaufteilung und militärischer Bereich

# Datentyphierarchie

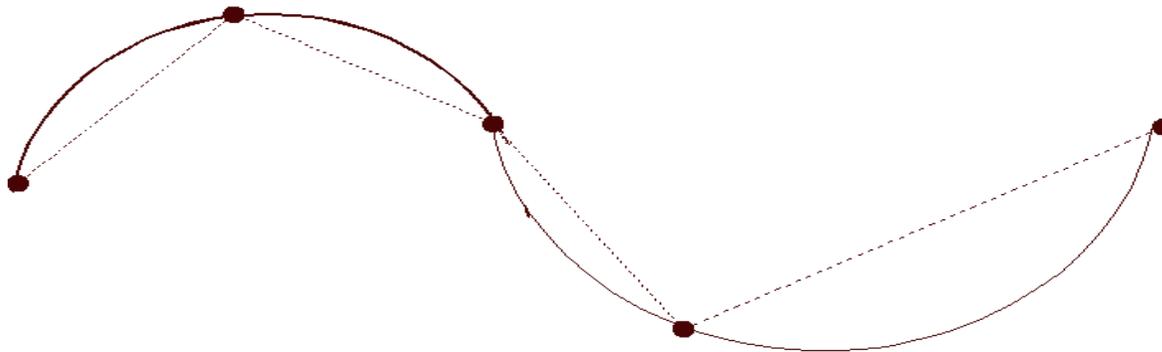


## ST\_Geometry

- Basistyp für alle anderen Typen
- stellt Standardmethoden bereit (Dimension, Typ,...)
- weiterführende Methoden zwischen ST\_Geometry-Werten, wie
  - ST\_Intersection
  - ST\_Union
  - ST\_Difference
  - ST\_SymDifference
  - ST\_Distance

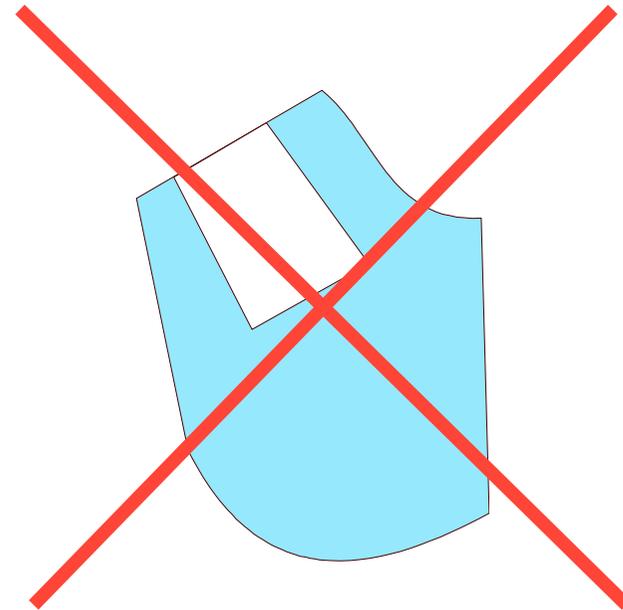
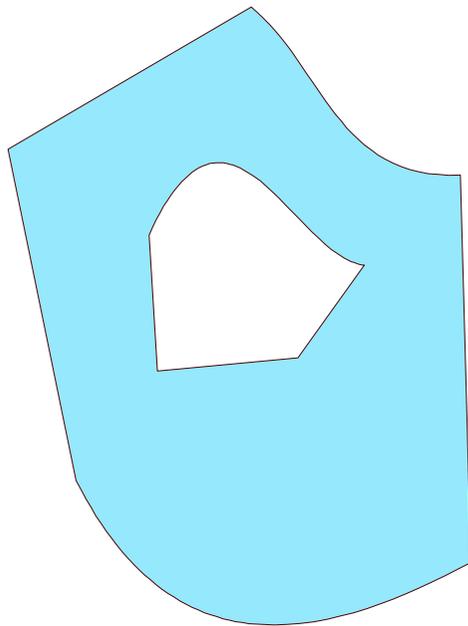
## ST\_Curve

- eindimensionale Geometrie
- Folge von Punkten
- Unterscheidung der Untertypen (ST\_LineString, ST\_CircularString)
- ST\_CompoundCurve



## ST\_CurvePolygon

- stellen Flächen dar
- aus Kurven zusammengesetzt



## Testen von Lagebeziehungen

- Testen auf Überlappung, Disjunktheit, Überschneidung, Berührungen, ... von zwei Geometrien
- die Methoden zum Testen dieser Beziehungen lassen sich auf eine Methode ST\_Relate zurückführen

a.ST\_Relate(b, `TF\*012\*\*\*`)

	Interior	Boundary	Exterior
Interior	1: $(\text{Interior}(a) \cap \text{Interior}(b)).$ ST_Dimension()	2: $(\text{Interior}(a) \cap \text{Boundary}(b)).$ ST_Dimension()	3: $(\text{Interior}(a) \cap \text{Exterior}(b)).$ ST_Dimension()
Boundary	4: $(\text{Boundary}(a) \cap \text{Interior}(b)).$ ST_Dimension()	5: $(\text{Boundary}(a) \cap \text{Boundary}(b)).$ ST_Dimension()	6: $(\text{Boundary}(a) \cap \text{Exterior}(b)).$ ST_Dimension()
Exterior	7: $(\text{Exterior}(a) \cap \text{Interior}(b)).$ ST_Dimension()	8: $(\text{Exterior}(a) \cap \text{Boundary}(b)).$ ST_Dimension()	9: $(\text{Exterior}(a) \cap \text{Exterior}(b)).$ ST_Dimension()

## ST\_Disjoint

$a.ST\_Disjoint(b) \Leftrightarrow$

$((Interior(a) \cap Interior(b) = \emptyset) \wedge$

$(Interior(a) \cap Boundary(b) = \emptyset) \wedge$

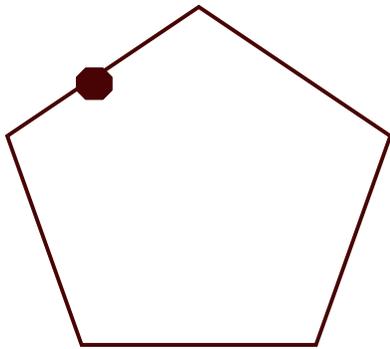
$(Boundary(a) \cap Interior(b) = \emptyset) \wedge$

$(Boundary(a) \cap Boundary(b) = \emptyset)) \Leftrightarrow$

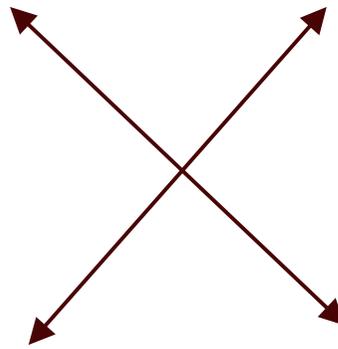
$a.ST\_Relate(b, \text{'FF*FF*FF*FF'})$

Auf ähnliche Weise könne auch andere Lagebeziehungen getestet werden.

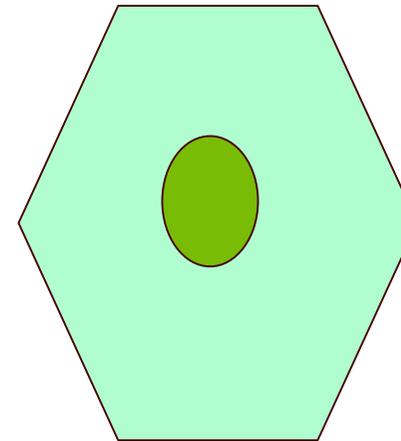
## Beispiele für Lagebeziehungen



Berührung  
ST\_Polygon/ST\_Point



Überschneidung  
ST\_LineString/ST\_LineString



Beinhalten  
ST\_Polygon/ST\_CurvePolygon

## externe Datenformate

zum Ein- bzw. Auslesen der Daten gibt es drei Formate

- Binärrepräsentation
- Textrepräsentation
- XML-Repräsentation (GML)

```
<Point gid="P1">
```

```
  <coord><X>5.0</X><Y>0.57</Y></coord>
```

```
</Point>
```

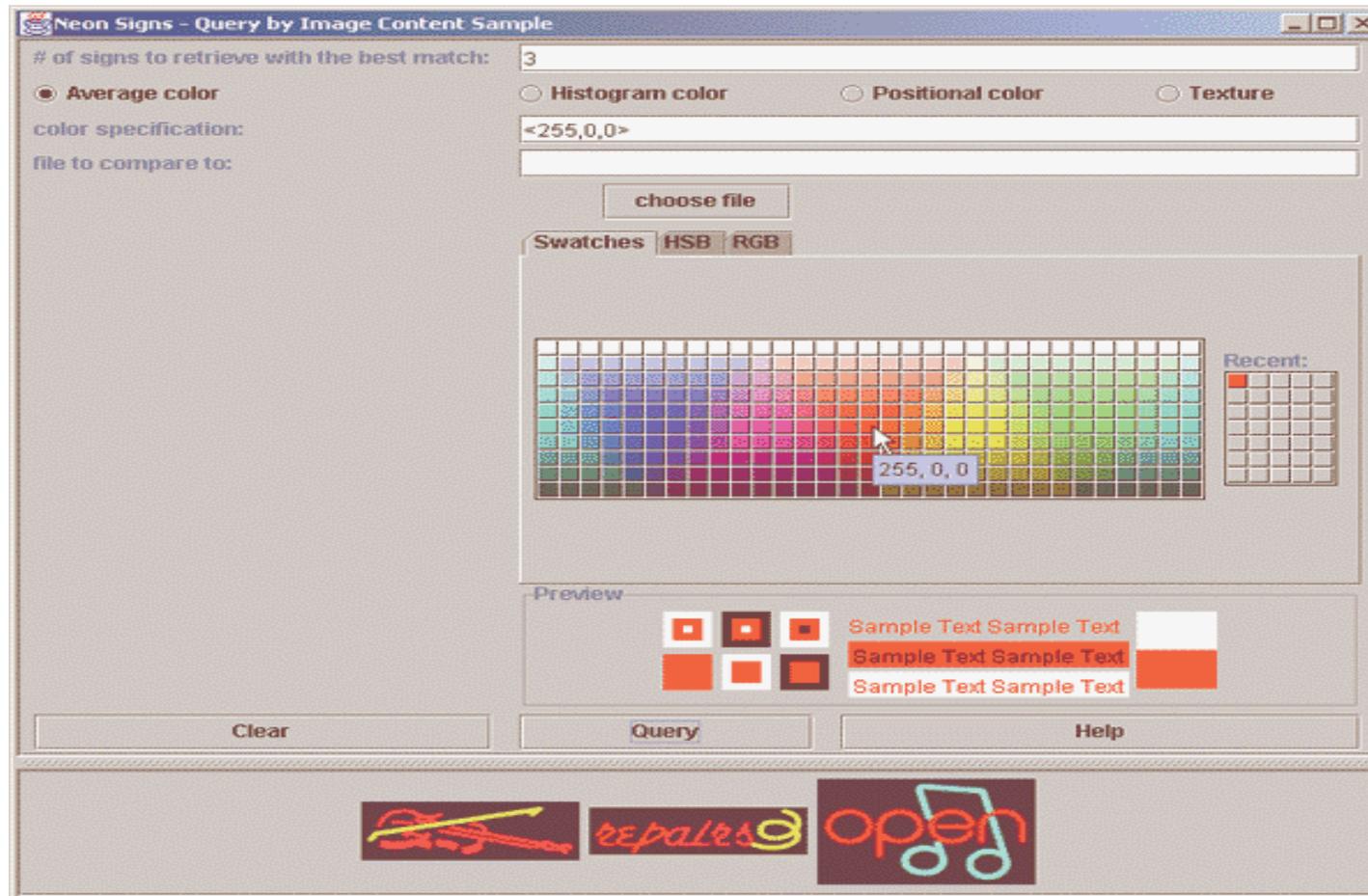
## Still Image: Einleitung

- zur Speicherung und Verarbeitung von Bildern in Datenbanken
- Suche nach Bildern mit Hilfe von „visuellen“ Prädikaten
- dazu Definition von Bildmerkmalen:
  - SI\_AverageColor
  - SI\_ColorHistogram
  - SI\_PositionalColor
  - SI\_Texture
  - SI\_FeatureList: Kombination

## SI\_AverageColor

- durchschnittliche Farbe
- wird berechnet durch:
  - Aufsummierung der einzelnen Farbwerte jedes Pixels (z.B. rot, grün, blau)
  - Division durch Anzahl der Pixel
- unterschiedliche Bilder können gleiche Werte besitzen

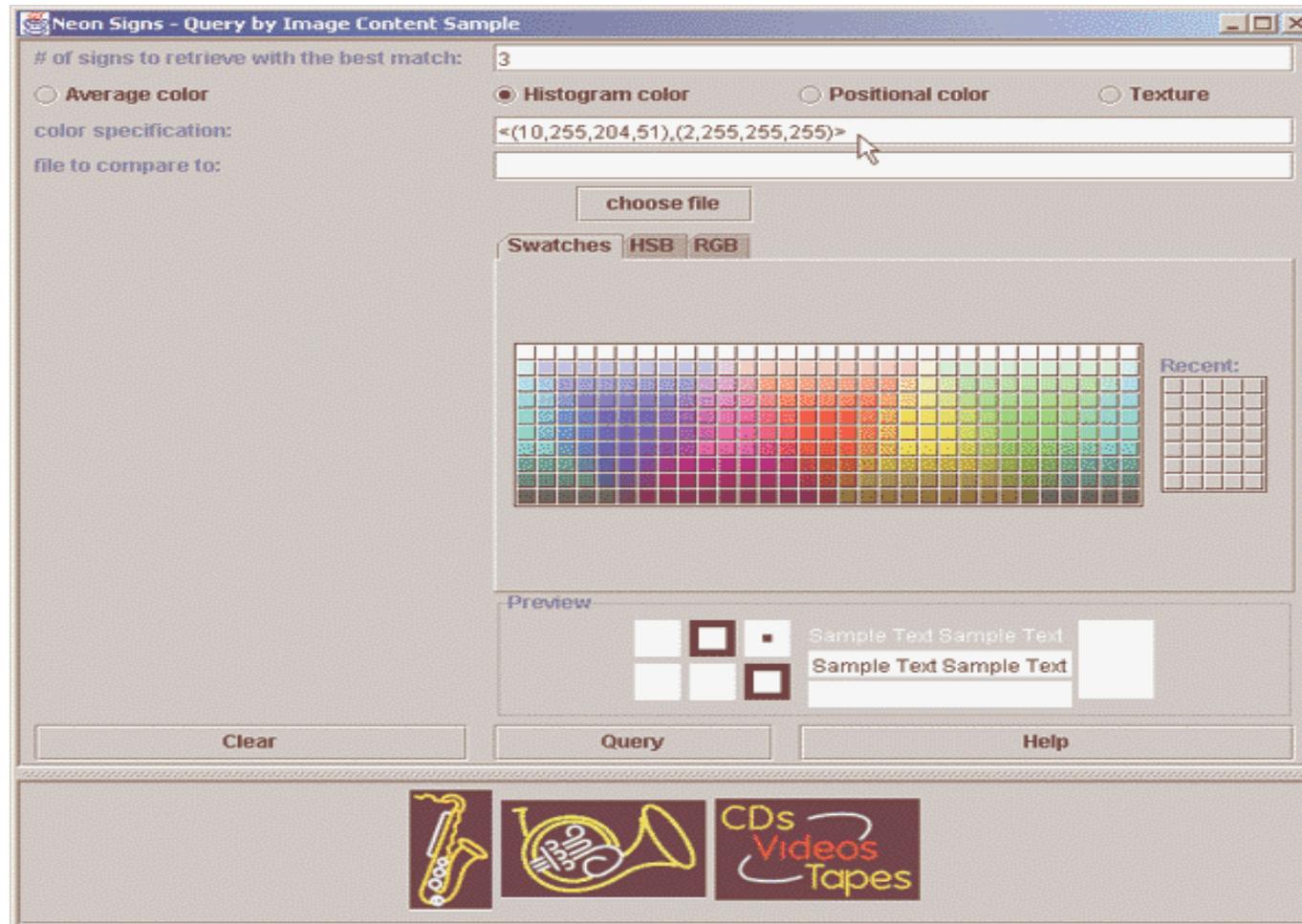
## SI\_AverageColor



## SI\_ColorHistogram

- relative Frequenz der Farben
- Zerlegung des Farbraumes
- zu jedem Teil des Farbraums werden die Pixel gezählt
- Division durch Anzahl der Pixel

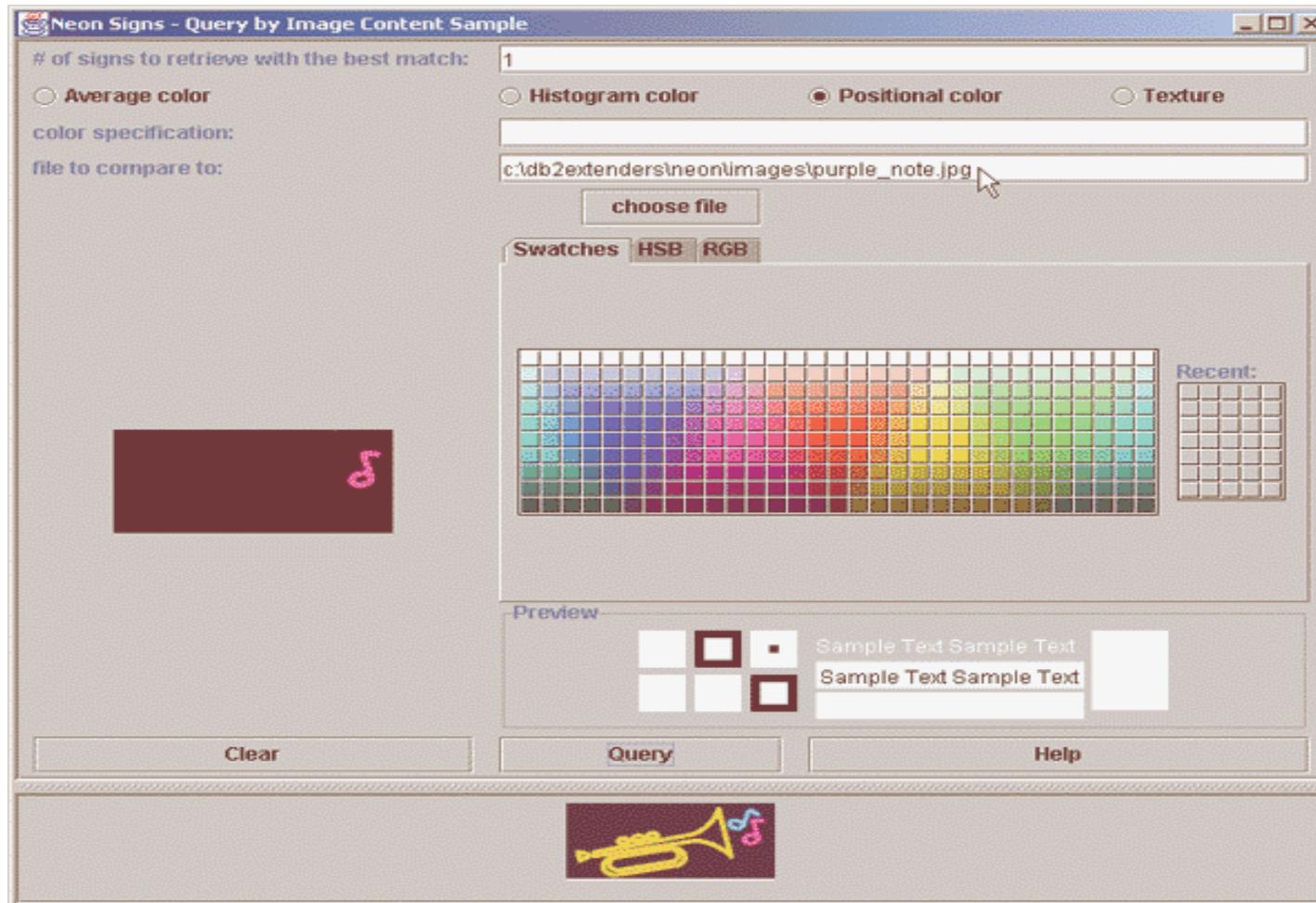
# SI\_ColorHistogram



## SI\_PositionalColor

- durchschnittliche Farbe für Teil des Bildes
- Bild wird in n mal m Rechtecke eingeteilt
- Berechnung für diese Bereiche analog zu SI\_AverageColor

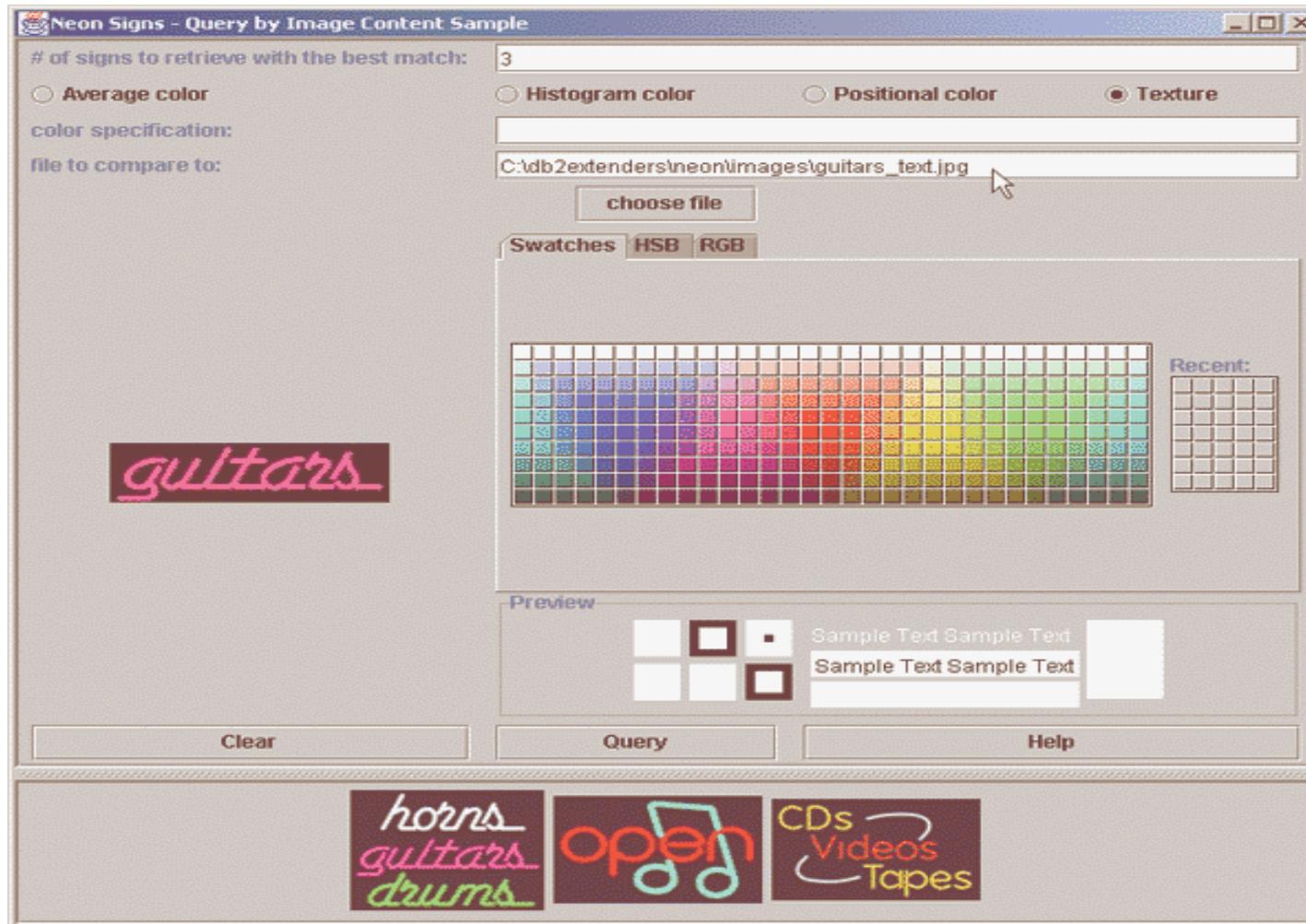
## SI\_PositionalColor



## SI\_Texture

- Charakterisierung anhand Grobheit und Kontrast
- keine genaueren Angaben im Standard
- große Variationen möglich

# SI\_Texture



## SQL/MM: Zusammenfassung

- neue Datentypen für Full Text, Spatial und Still Image
- bessere vereinheitlichte Möglichkeiten zur Suche in Texten, geographischen Daten und Bildern
- Relevanz in kommerziellen DBS

Ende