2. Mở Rộng

Trong tài liệu này, có lẽ sẽ thấy được nhiều hệ thống và phương pháp truy xuất hình ảnh dựa trên nội dung. Trong phần này, chúng ta thảo luận về các phần chức nâng, thư mục hình ảnh, sự ảnh hưởng của màn hình (cảnh) và vai trò kiến thức tên miền, và khoảng cách ngữ nghĩa giữ những đặc điểm hình ảnh và người dùng.

2.1 Chức năng của phương pháp truy xuất hình ảnh dựa trên nội dung.

Trong [31], chúng ta thấy được 3 loại mục đích sử dụng khi dùng hệ thống, nhìn hình 3

* Có một lớp đại cương những phương pháp và những hệ thống về mục đích trình duyệt thông qua bộ ảnh lớn từ những nguồn không rõ. Những người dùng của Nghiên *Cứu theo Hiệp hội* lúc bắt đầu không có mục đích gì dặc biệt gì khác hơn là tìm kiếm gì đó thú vị. Nghiên cứu theo hiệp hội thường ngụ ý lặp lại sàng lọc của việc tìm kiếm, những sự đồng dạng và ví dụ đã được bắt đầu nghiên cứu. Những hệ thống trong loại này thường tương tác rất cao, nơi mà thông số kỹ thuật có lẽ được phát thảo [30] hoặc được ví dụ bởi hình ảnh. Ví dụ thực tế cũ nhất của một hệ thống có lẽ là [88]. Kết quả của nghiên cứu có thể thao tác tác động với nhau bởi sự phản hồi liên quan [68], [51]. Để hỗ trợ những yêu cầu cho kết quả liên quan (xác đáng), thì những nguồn khác hơn cả hình ảnh được sử dụng, ví dụ như [168], [21].
* Một lớp người dùng khác đặt mục tiêu nghiên cứu ở những hình ảnh chi tiết. Nghiên cứu có lẽ vì một bản sao rõ ràng theo hình ảnh trong trí nhớ, như trong nghiên cứu catalogues nghệ thuật, ví dụ [48]. Mục tiêu nghiên cứu có lẽ cũng bởi vì một hình ảnh khác của cùng một vật thể mà người dùng đã có hình sẵn. Đây là một mục tiêu nghiên cứu bởi vật mẫu. Mục tiêu nghiên cứu cũng có lẽ được ứng dụng khi người dùng có một hình ảnh đặc biệt trong ký ức và mục tiêu được định rõ một cách có tương tác lẫn nhau tương tự như một nhóm ví dụ đã cho, ví dụ [31]. Nhìn chung, những hệ thống này phù hợp với những nghiên cứu như tem, mỹ thuật, linh kiện (thành phần công nghiệp) và catalogue.
* Lớp thứ ba của những ứng dụng, nghiên cứu theo loại, mục đích để truy lại một hình ảnh bất kỳ tượng trung cho một lớp chi tiết. Nó có thể là một trường hộp mà người dùng có một vật mẫu và nghiên cứu về yếu tố cùng một lớp. Nhiều loại có thể được xuất phát từ những ký hiệu (nhãn: label) hoặc nổi lên từ cơ sở dữ liệu [170], [186]. Trong nghiên cứu theo loại, người dùng có thể có sẵn một nhóm hình ảnh và nghiên cứu về những hình ảnh bổ sung cùng lớp [28]. Một ứng dụng đặc trưng của nghiên cứu theo loại là catologue đa dạng. Theo [74], [79] những hệ thống được thiết kế để phân loại nhãn hiệu. Những hệ thống trong loại này thường tương tác với một tên miền đặc biệt định nghĩa theo sự tương đồng.

Ba loại sử dụng này không là tất cả bài báo cáo [42]. Một nghiên cứu [121] của nhà báo đã định nghĩa năm dạng sử dụng: nghiên cứu vì một hình ảnh chi tiết (đặc biệt), truy cập thông thường để tạo một sự lựa chọn tương tác, nghiên cứu vì một hình ảnh đi với một bài báo bảng biểu, nghiên cứu để minh họa dữ liệu, và nghiên cứu vì thay thế chỉ giá trị mỹ quan của hình ảnh. Một mục đích để đề ra một sự phân nhóm chung theo yêu cầu của người dùng vì hình ảnh tĩnh và động được tìm thấy ở [6]. Điều này và những nghiên cứu tương tự phát hiện ra rằng một dãy truy vấn thì rộng hơn chỉ truy hình ảnh dựa trên sự xuất hiện hay không xuất hiện của vật thể theo các đặc điểm hình ảnh đơn giản.

2.2 Tên miền hình ảnh và khoảng cách nhận cảm.

Trong thư mục hình ảnh dưới sự cân nhắc – Miền hình ảnh I – Có một sự phân biệt theo từng bậc giữa những tên miền hẹp và rộng [160]. Ở cuối một phổ, chúng ta có tên miền hẹp:

*Một tên miền hẹp có tính biến thiên giới hạn và có thể đoán trước được trong tất cả những khía cạnh liên quan của sự xuất hiện của nó.*

Trong một tên miền, bất cứ ai tìm ra một tính biến thiên giới hạn của nội dung hình ảnh. Thường thì chụp quay chi tiết cũng tương tự trên tất cả tên miền. Trong tên miền hẹp của lithographs, ví dụ, sự chụp quay dưới đèn trắng với cảnh chính diện và không có tình trạng bị bít. Cũng là khi sự xuất hiện diện của vật thể có tính biến thiên giới hạn, sự mô tả ngữ nghĩa của hình ảnh thường dễ xác nhận bởi rộng và nhất quán. Một ví dụ khác của một tên miền hẹp là một bố cục cảnh mặt chính diện được chụp quay chống đối với phông nền sáng (rõ ràng). Mặc dù mỗi khuôn mặt là độc nhất và có tính biến thiên rộng ở những chi tiết hình ảnh. Có những liên kết rõ ràng về hình học, vật lý và màu sắc liên quan điều khiển tên miền. Tên miền sẽ rộng hơn khi khuôn mặt được chụp từ cảnh đông đúc hoặc từ cảnh ngoài trời. Trong trường hợp đó, sự biến thiên của sự chiếu sáng, sự lộn xộn của phông cảnh, tình trạng bị bít, và điểm nhìn sẽ có ảnh hưởng cơ bản lên sự phân tích.

Ở điểm cuối khác của một phổ, chúng ta có tên miền rộng:

*Tên miền rộng có tính biến thiên không giới hạn và không thể đoán trước được trong tất cả sự hiện diện của nó thậm trí cùng ngữ nghĩa.*

Ở những tên miền rộng, hình ảnh thì đa nghĩa và ngữ nghĩa học của nó được mô tả chỉ từng phần. Nó có lẽ là trường hợp có những vật thể dễ thấy trong phông cảnh vì lớp vật thể thì không được biết hoặc thậm trí sự thể hiện của phồng nền không là duy nhất. Một lớp rộng của hình ảnh có thể được tìm trong kho hình ảnh rộng [168] hoặc kho trữ hình ảnh khác [42]. Lớp rộng nhất có sẵn đến ngày thì bố cục hình ảnh có sẵn trên mạng.

Nhiều vấn đề quan tâm thực tiễn có một tên miền hình ảnh giữa những điểm cuối cực hạn của phổ, nhìn hình 4. Khái niệm rộng và hẹp giúp mô tả đặc điểm nhiều dạng sử dụng, trong việc chọn lựa đặc điểm và trong thiết kế hệ thống. Trong một tên miền ảnh rộng, khoảng cách giữa mô tả đặc điểm và diễn dịch ngữ nghĩa thì thường rộng. Đối với miền hẹp, tên miền hình ảnh chuyên môn hóa thì khoảng cách giữa mô tả đặc điểm với diễn dịch ngữ nghĩa thường nhỏ hơn, vì thế tên miên – dạng đặc biệt có lẽ giúp được. Đối với khuôn mặt, nhiều dạng hình học được đề xuất cũng như thống kê đã nghiên cứu các dạng [127]. Những dạng có sử dụng máy điện toán này không có sẵn cho tên miền hình ảnh rộng bởi vì dạng có sử dụng máy điện tính này yêu cầu số lượng rất lợn.

Đối với tên miền hình ảnh rộng đặc thù, bất cứ ai phải sử dụng đến nguyên tắc có hiệu lực một cách thông thường. Có phải sự chiếu sáng tên miền là đèn màu trắng hay đèn có màu? Có phải nó giả định vật thể được xác định và có thể thấy được một cách toàn diện hay không? Có phải chụp quay 2D phông nền 2D hoặc chụp quay 2D phông nền 3D? Đặc trưng cụ thể của độ chiếu sáng, hiện diện hoặc không hiện diện của sự bít kín, sự hỗn loạn và sự khác nhau từ điểm nhìn của máy ảnh xác định nhu cầu dựa trên phương pháp truy xuất.

Khoảng cách nhận cảm là khoảng cách giữ vật thể trong thế giới và thông tin mô tả (có sử dụng máy tính điện toán) nhận được từ chụp quay phông nền đó.

Khoảng cách nhận tạo tạo nên sự mô tả vật thể một vấn đề không đúng đắn: nó tiến hành tình trạng không chắc chắn trong cái mà không được biết về trạng thái của vật thể. Khoảng cách nhận cảm được poignant đặc thù (một cách đặc biết) khi điều kiện chụp quay nhận biết rõ ràng bị chệch. Sự Chụp quay 2D khác với vật thể 3D có thể đồng nhất. Không có nhận biết hơn nữa, bất kì ai cũng phải quyết định rằng họ có thể miêu tả (trình bày) vật thể. Cũng như chụp quay 2D của phông cảnh 3D bao gồm thông tin đồng nhất cho phông nên đó và điều đó có nghĩa nhưng không phải bất kì ai cũng phần nào của thông tin liên quan đến phông cảnh. Tình trạng không chắc chắn bởi vì khoảng cách nhận cảm không chỉ giữ cho điểm cảnh mà còn cho sự tình trạng bí kín (nơi mà phần cần thiết cho rằng 2 phần vật thể có thể ra khỏi tầm nhìn), sự hỗn loạn và độ chiếu sáng.

So với sự diễn dịch đan xen có thể làm giảm đi khoảng cách nhận cảm. Hệ thống truy xuất hình ảnh dựa trên nội dung có thể cung cấp hỗ trợ trong định hướng này thông qua sự loại bỏ giữ một vài giải thích tiềm ẩn, nhiều như là trong xử lý ngôn ngữ tự nhiên.

2.3 Kiến thức tên miền. (Nhận biết tên miền)

Trong nghiên cứu hình ảnh, sự trình bày rõ ràng của nhận biết tên miền thì rất quan trọng để làm giảm khoảng cách nhận cảm. Giữa các nguồn nhận biết chung, we tôi đề cập đến:

* Những định luật cú pháp (bằng chữ/hằng số) đẳng thức và sự đồng dạng xác định quan hệ giữa chất lượng hình ảnh hoặc đặc điểm hình ảnh bất chấp tính vật lý hoặc những nguyên nhân giác quan. Ví dụ như, xem xét hai hình ảnh tương đồng bởi vì cả hai phô bày vài bóng xanh dương được chọn phần phía trên của nó được tạo ra trong sự tách phông cảnh ngoài trời từ hình ảnh khác. Đó là cú pháp sự đồng dạng bởi vì phương pháp không tạo ra sự tham chiếu đến lí do bởi sự đồng dạng này có tồn tại (trong trường hợp này là sự phân tán của bầu trời) hoặc đến những lý do giác quan bởi hai hình ảnh sẽ xuất hiện tương đồng đến người quan sát. Bởi cùng một chìa khóa số, khoảng màu RGB có hiệu quả trong hằng số đồng dạng (như hiệu quả trong nghệ thuật [65]) trong khi đó không thể hiện quá trình cấu tạo màu sắc vật lý hoặc quá trình cấu tạo màu sắc cảm giác.
* Định luật mô tả đẳng thức và sự đồng dạng cảm giác con người thì rất quan trọng bởi vì nó xác định đẳng thức giống nhau dựa trên đẳng thức kinh nghiệm người dùng. Trong màu sắc, CIE – Lab và Munsell-spaces đã được thiết kế để thích hợp với cảm giác con người về sự đồng dạng màu sắc. Nếu sự đánh giá của người quan sát vật thể được dựa trên giác quan của mục chắc chắn dễ thấy trên hình ảnh [177], đó là tính tự nhiên đến sự tính toán trực tiếp những đặc điểm tên miền rộng đến những điểm và khu vị này [157], [138]. Cũng giống như, cấu trúc về mặt sinh học đáng tin cậy [76] của đơn vị xử lý xung quanh trung tâm có thể chọn lựa khu vực mà con người sẽ tập trung vào đầu tiên.
* Định luật vật lý mô tả đẳng thức và sự khác nhau giữa hình ảnh dưới sự khác nhau trong phông nền và tính chất bề mặt vật thể. Tính vật lý trong độ chiếu sáng, độ tương phản bề mặt, và cấu tạo hình ảnh có một hiệu quả chung trên hình ảnh. Định luật chung của vật lý có lẽ được dùng cho những lớp rộng của vật thể. Một ví dụ thông thường là định luật cho sự tương phản ánh sáng không đổi làm mờ đi vật thể. Những định luật này được lợi dụng để thiết kế đặc điểm màu sắc cái mà biểu thị bất đẳng thức của kiểu dáng và điểm ảnh.
* Quy tắc hình học và tôpô miêu tả đẳng thức và sự khác nhau giữ khuôn mẫu trong không gian. Khi hai vật thể trong hình học bằng thì tính chất vật lý bề mặt của chúng hoặc điều kiện vật lý của phông nền có thể khác nhau. Như một ví dụ trong quy tắc hình học được dùng trong truy xuất vì tất cả hình ảnh với độ sau và chi tiết cục bộ gần đường nằm ngang sẽ xuất hiện nhỏ hơn. Cũng như đường nằm ngang rõ ràng là hình học vì một đường ảo bao gồm những tiêu điểm. Một ví dụ khác của định luật hình học là sự trình bày không gian [22] hoặc những quan hệ tôpô [172] giữa hai vật thể.
* Quy tắt loại cơ bản mã hóa đặc điểm chung đến lớp z trong không gian của tất cả khái niệm Z. Nếu z là lớp của tất cả ấm trà (teapots), những đặc đểm bao gồm sự hiện diện của cổ bình. Nhiều loài hầu như được sử dụng dành riêng trong một tên miền hẹp. Nhận biết tên miền có thể theo hình thức thêm ràng buộc đến chất lượng hình ảnh hằng số, tính vật lý bổ sung hoặc quy luật hình học, hoặc tùy chỉnh nhân tạo tên miền đặc trưng. Khi tên miền đang vạch ra bản vẽ, nhận biết hình học chi tiết sẽ đi theo hướng phát hiện ra các ký hiệu. Trong nghệ thuật thời cổ xưa, màu sắc và vị trí tương đối của vật thế có ý nghĩa ký hiệu [30], phát ra một cài đặt hữu ích bắt buộc trong nghiên cứu. Mỗi chức năng tên miền có một cài đặt bảo bật bắt buộc.
* Cuối cùng, tùy chỉnh theo nhân tạo hoặc mẫu liên quan nhân tạo mở đầu quy luật đăng thức và sự khác nhau cơ sở văn hóa. Dưới nền văn hóa, chúng ta cũng giả định ngôn ngữ. Trong nghiên cứu bức hình trong nhà, bất kì ai cũng có thể kiểm tra nhiều đường thẳng và những gốc vuông theo tiêu chuẩn lựa chọn đầu tiên. Những dụng cụ có một kích thước nhất định để cho phép bắt sóng. Thời xác định màu sắc [95].

Những định luật này được sắp xếp như hình 5

2.4 Cách dùng và người dùng, khoảng cách ngữ nghĩa.

Chúng ta cho rằng hầu hết những sự thất vọng với hệ thống truy xuất ban đầu đến từ việc thiếu hụt tổ chức sự tồn tại khoảng cách ngữ nghĩa học và hệ quả của nó cho cài đặt hệ thống.

*Khoảng cách ngữ nghĩa học là thiếu hụt sự trùng hợp giữa thông tin mà bất cứ ai có thể trích từ dữ liệu hình ảnh và sự diễn dịch mà dữ liệu giống nhau có cho người dùng trong tình huống đã cho sẵn.*

Mô tả ngôn ngữ là hầu như luôn thuộc văn cảnh trong khi một hình ảnh có thể tồn tại bởi chính nó. Mô tả ngôn ngữ một hình ảnh là một sự nản chí, có lẽ không thể tác vụ [146]. Một người dùng tìm kiếm hình ảnh bao gồm những vật thể nhất định hoặc chuyển một tin nhất nhất định. Những mô tả hình ảnh theo mặt khác, tin vào đặc điểm điều khiển bằng dự liệu và cả hai có thể không được kết nối. Hiệp hội hệ thống ngữ nghĩa hoàn thành với dữ liệu hình ảnh sẽ kế thừa ít nhất để giải quyết vấn đề nhận biết vật thể chung từ một hình ảnh đơn. Vì vấn đề này chưa được giải quyết, nghiên cứu được tập trung vào nhiều phương pháp khác để kết hợp những ngữ nghĩa trình độ cao hơn với điều khiển dự liệu quan sát được.

Như được thể hiện ở hình 2, những nghĩa trực tiếp nhất của đặc điểm ngữ nghĩa đòi hỏi sự chú thích bởi những từ khóa hoặc những thuyết minh. Điều này làm giảm truy cập dựa trên nội dung đến thông tin truy xuất [135]. Những vật thể thông thường đến nhãn hiệu thực tế đều đắc và phủ sóng. Về mặt chi phí, hàng ngàn nhãn hiệu của hình ảnh là một công việc ngổn ngang và đắc giá tạo nên cấp bậc mà sự triển khai của sự cân bằng kinh tế đằng sau cơ sở dự liệu giống như giảm xuống. Để giải quyết vấn đề này, hệ thống trong [21], [142] dùng một chương trình nghiên cứu mạng, thu thập hình ảnh và chèn chúng vào nguyên tắc phân loại được xác định trước trên cơ sở của câu lệch xung quanh chúng. Một phương thức tiếp cận tương tự cho thư viện số được thực hiện bởi [24]. Về mặt phủ sóng, nhãn hiệu thì hiếm khi thực hiện, ngữ cảnh nhạy cảm, và trong trường hợp này, có một dấu hiệu yêu cầu phân số cái mà ngữ nghĩa học không thể bị chụp bởi nhãn hiệu đơn [6], [64]. Cả hai phương pháp sẽ che đẩy khoảng cách ngữ nghĩa chỉ trong những trường hợp bị cô lập.

2.5 Sự thảo luận mở rộng.

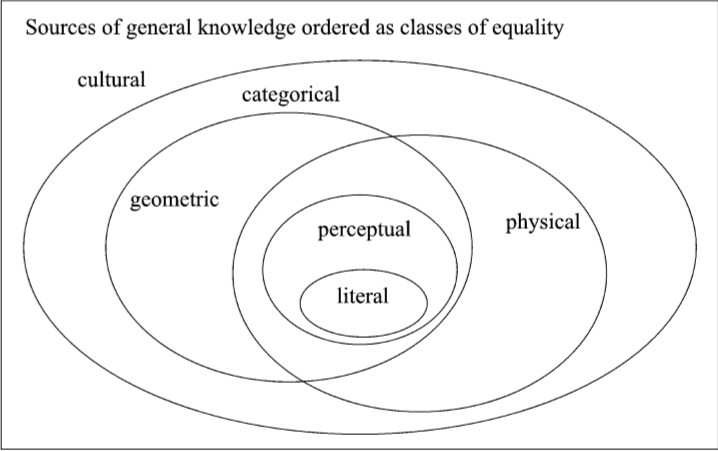
Điểm cốt lõi trong truy xuất dựa trên nội dung là người dùng tìm kiếm sự đồng dạng ngữ nghĩa học nhưng cơ sở dữ liệu có thể chỉ cung cấp sự đồng dạng bởi xử lý dự liệu. Đây là điều chúng ta đã gọi là khoảng cách ngữ nghĩa. Vào cùng thời điểm, khoảng cách nhận cảm giữa đặc tính trong một hình ảnh và đặc tính trong vật thể đóng một vai trò giới hạn trong truy xuất nội dung của hình ảnh.

Chúng ta đã thảo luận ứng dụng của truy xuất dựa trên nội dung trong 3 loại mở rộng: nghiên cứu theo mục tiêu, nghiên cứu theo phạm trù và nghiên cứu bởi hiệp hội (liên kết). Nghiên cứu theo mục tiêu kết nối với truyền thống của khuôn mẫu kết nối với tầm nhìn máy tính. Những thử thách mới trong truy xuất dựa trên nội dung là số lượng lớn vật thể trong nghiên cứu, sự đặc tả truy vấn không đầy đủ, mô tả hình ảnh không đầy đủ, và điều kiện phông nền và vị trí vật thể chuyển động. Nghiên cứu theo phạm trù xây sự nhận dạng vật thể và phương pháp nhận dạng khuôn mẫu thống kê trong tầm nhìn máy tính. Những thách thức trong truy xuất dựa trên nội dung được so sánh để hoàn thành sự nhận viết vật thể là kết quả thủ thuật có tính tương tác, thường là số lượng lớn nhiều lớp vật thể, và sự hiện diện của thời kỳ đào tạo rõ ràng cho đặc điểm và sự điều chỉnh phân loại.

Nghiên cứu theo liên kết thì khác biệt xa hơn với hầu hết tầm nhìn truyền thống máy tính. Nó bị cản trở bởi hầu như khoảng cách ngữ nghĩa. Chỉ cần khoảng cách ở đây, sử dụng truy xuất dựa trên nội dung cho trình duyệt sẽ không trong quyền hạn của công chúng bởi vì con người được quen với việc tin cậy dấu vết ngữ nghĩa trực tiếp khi họ thấy một hình ảnh. Mục đích của hệ thống truy xuất dựa trên nội dung cần được cung cấp sự hỗ trợ lớn nhất trong cầu nối khoảng cách ngữ nghĩa giữa tính đơn giản của đặc điểm hình ảnh có sẵn và sự giàu có của người dùng ngữ nghĩa.

Chúng ta phân tích những đặc tính của tên miền hình ảnh, nhận biết tên miền và loại sử dụng vì những yếu tố cơ bản xác định chức năng của một hệ thống. Một sự đặc thù quan trọng là giữ những tên miền rộng và hẹp. Tên miền càng rộng, trình duyệt càng nhiều hoặc nghiên cứu bởi liên hết có thể là phương pháp đúng. Tên miền càng hẹp, ứng dụng tương tự của nhận biết tên miền sẽ thành công. *Thách thức cho cơ cấu nghiên cứu hình ảnh trên tên miền rộng là để điều chỉnh cơ cấu tên miền hẹp mà người dùng quan tâm đến thông qua chi tiết cụ thể, những ví dụ và sự tương tác.*

|  |  |  |
| --- | --- | --- |
| Tên miền hẹp và rộng trong truy xuất hình ảnh | | |
|  | Hẹp | Rộng |
| Sự khác biệt nội dung | Thấp | Cao |
| Nguồn nhận biết | Chi tiết | Theo loại |
| Ngữ Nghĩa học | Đồng nhất | Không đồng nhất |
| Sự thật nền tảng | Có thể đúng | Không thể đúng |
| Nội dung mô tả | Khách quan | Chủ quan |
| Phông nền và nhận cảm | Có thể điều khiển được | Không chắc |
| Ứng dụng mục tiêu | Chi tiết | Theo loại |
| Loại ứng dụng | Chuyên gia | Công chúng |
| Công cụ | Định hướng mô hình  bất biến chi tiết | Giác quan, văn hóa,  Bất biến theo loại |
| Tương tác | Giới hạn | Tràn lan, tương tác lẫn nhau |
| Đánh giá | Số lượng | Chất lượng |
| Hệ thống kiến trúc | Định hướng cơ sở dự liệu biến đổi | Định hướng cơ sở dữ liệu mô đun |
| Kích thước | Trung bình | Rộng hoặc rất rộng |
| Nguồn cảm hứng | Sự nhận biết vật thể | Truy xuất thông tin |

*Hình 4: bảng kham khảo sơ lược tên miền hẹp và rộng*

**Nguồn nhận biết chung sắp xếp theo lớp đẳng thức**

Hình 5: Khi tìm kiếm kiếm theo từ khóa “ghế”, chúng tôi có lẽ được thỏa mãn với bất kì vật thể nào dưới cái tên đó, đó là chúng tôi tìm kiếm theo đẳng thức định nghĩa bởi con người. Khi chúng tôi tìm kiếm theo cái tên những ghế một chân dài, chúng tôi thêm một lớp ràng buộc bổ sung vào lớp chung và hạn chế lớp đẳng thức. Sự nắm giữ được giống nhau khi tìm kiếm cho từ khóa “ghế đỏ”, thêm một điều kiện độc lập hạn chế hình học. Khi chúng tôi tìm kiếm một chiếc ghế tương đương với một chiếc ghế đã định, ít nhất phải có chất lượng vật lý và hình học. Cuối cùng, khi chúng tôi tìm kiếm chính xác cùng một hình ảnh của chiếc ghế đó, đẳng thức theo nghĩa đen được yêu cầu, tất nhiên vẫn bỏ qua các biến thể do tiếng ồn.