An Efficient Global Optimization Algorithm for Nonlinear Sum-of-Ratios Problem *†

Yun-Chol Jong[a]

[a] Center of Natural Sciences, University of Science, Pyongyang, DPR Korea

May 3, 2012

Abstract

This paper presents a practical method for finding the globally optimal solution to nonlinear sum-of-ratios problem arising in image processing, engineering and management. Unlike traditional methods which may get trapped in local minima due to the non-convex nature of this problem, our approach provides a theoretical guarantee of global optimality. Our algorithm is based on solving a sequence of convex programming problems and has global linear and local superlinear/quadratic rate of convergence. The practical efficiency of the algorithm is demonstrated by numerical experiments for synthetic data.

2010 Mathematics subject Classification: 90C26, 90C32, 65K05

Keywords: Fractional programming, non-convex optimization, global optimization algorithm, sum-of ratios problem, guaranteed global optimality.

1 Introduction

The sum-of-ratios problem, which is to minimize (maximize) a sum of several fractional functions subject to convex constraints, is a non-convex optimization problem that is difficult to solve by traditional optimization methods. The problem arises in many applications such as optimization of the average element shape quality in the finite element method, computer

*E-mail: yuncholjong@yahoo.com

†Address: Gwahak-1 dong, Unjong distrct, Pyongyang, DPR Korea

graphics and management ([2],[3],[4]). In [4], many problems of projective geometry including multiview triangulation, camera resectioning and homography estimation have been formulated as the sum-of-ratios problem and a branch-and-bound method has been proposed to find its global solution which relies on recent developments in fractional programming and the theory of convex underestimators ([6],[7],[8],[9],[14],[15]). In the method of [4], number of variables increases as twice as the number of fractional functions involving in the sum and a second-order cone programming problem is needed to solve for obtaining a lower bound of the optimal value in each iteration. Their algorithm is provably optimal, that is, given any tolerance ϵ, if the optimization problem is feasible, the algorithm returns a solution which is at most ϵ far from the global optimum. The branch-and-bound method requires a lot of computations, has low convergence and it is not easy to find a reasonable branching strategy. Recently there has been some progress made towards finding the global solution to a few of these optimization problems ([16],[17]). However, the resulting algorithm is numerically unstable, computationally expensive and does not generalize for more views or harder problems like resectioning. In [18], linear matrix inequalities were used to approximate the global optimum, but no guarantee of actually obtaining the global optimum is given. Also, there are unsolved problems concerning numerical stability. Robustification using L_1-norm was presented in [19], but the approach is restricted to the affine camera model.

In this paper, an efficient global optimization algorithm is presented which transforms the sum-of-ratios problem into parametric convex programming problem and finds the global solution successfully.

2 Equivalent parametric convex programming

The sum-of-ratios problem seeks to minimize the sum of fractional functions subject to convex constraints, which is formulated as follows.

$$\begin{aligned} &\min F(x) = \sum_{i=1}^{N} F_i(x), \\ &\qquad\text{subject to} \quad g_i(x) \leq 0, i = 1, \cdots, m, \\ &\qquad x \in R^n \end{aligned} \tag{2.1}$$

where $F_i(x) = \frac{f_i(x)}{h_i(x)}, i = 1, \cdots, N$, and $f_i(x), g_i(x)$ and $-h_i(x)$ are twice continuously differentiable convex functions.

Let $X = \{x \in R^n | g_i(x) \leq 0, i = 1, \cdots, m\}$. It is assumed that $f_i(x) \geq 0$ and $h_i(x) > 0$ for every $x \in X$, and that $intX = \{x \in R^n | g_i(x) < 0, i = 1, \cdots, m\} \neq \emptyset$. Even with these restrictions the above problem is NP-complete [5].

It is easy to see that the problem (2.1) is equivalent to the following problem.

$$
\begin{aligned}
\min F(x) = \sum_{i=1}^{N} \beta_i,& \\
&\text{subject to} \quad F_i(x) \le \beta_i, i = 1, ..., N, \\
&g_i(x) \le 0, i = 1, ..., m, \\
&x \in R^n
\end{aligned}
\tag{2.2}
$$

Lemma 2.1. *If $(\bar{x}, \bar{\beta})$ is the solution of the problem (2.2), then there exist $\bar{u}_i, i = 1, \cdots, N$ such that $\bar{x}$ is a solution of the following problem for $u = \bar{u}$ and $\beta = \bar{\beta}$.*

$$
\begin{aligned}
\min \sum_{i=1}^{N} u_i(f_i(x) - \beta_i h_i(x)),& \\
&\textit{subject to} \quad g_i(x) \le 0, i = 1, ..., m, \\
&x \in R^n
\end{aligned}
\tag{2.3}
$$

And $\bar{x}$ also satisfies the following system of equations for $u = \bar{u}$ and $\beta = \bar{\beta}$:

$$u_i = \frac{1}{h_i(x)}, i = 1, ..., N \tag{2.4}$$

$$f_i(x) - \beta_i h_i(x) = 0, i = 1, ..., N \tag{2.5}$$

Proof. The constraint $F_i(x) \le \beta_i$ is equivalent to $f_i(x) - \beta_i h_i(x) \le 0$. Let's define the following function for the problem (2.2).

$$L(x, \beta, w, u, v) = w \sum_{i=1}^{N} \beta_i + \sum_{i=1}^{N} u_i(f_i(x) - \beta_i h_i(x)) + \sum_{i=1}^{m} v_i g_i(x).$$

By Fritz-John optimality condition (Theorem 4.2.8 of [1]), there exist $\bar{w}, \bar{u} = (\bar{u_1}, \cdots, \bar{u_N})$ and $\bar{v} = (\bar{v_1}, \cdots, \bar{v_m})$ such that

$$\frac{\partial L}{\partial x} = \sum_{i=1}^{N} \bar{u_i}(\nabla f_i(\bar{x}) - \bar{\beta_i} \nabla h_i(\bar{x})) + \sum_{i=1}^{m} \bar{v_i} \nabla g_i(\bar{x}) = 0 \tag{2.6}$$

$$\frac{\partial L}{\partial \beta_i} = \bar{w} - \bar{u_i} h_i(\bar{x}) = 0, i = 1, \cdots, N \tag{2.7}$$

$$\bar{u_i} \frac{\partial L}{\partial u_i} = \bar{u_i}(f_i(\bar{x}) - \bar{\beta_i} h_i(\bar{x})) = 0, i = 1, \cdots, N \tag{2.8}$$

$$v_i \frac{\partial L}{\partial v_i} = \bar{v_i} g_i(\bar{x}) = 0, i = 1, \cdots, m \tag{2.9}$$

$$g_i(\bar{x}) \leq 0, \bar{v}_i \geq 0, i = 1, \cdots, m \tag{2.10}$$

$$f_i(\bar{x}) - \bar{\beta}_i h_i(\bar{x}) \leq 0, \bar{u}_i \geq 0, i = 1, \cdots, N \tag{2.11}$$

$$\bar{w} \geq 0, (\bar{w}, \bar{u}, \bar{v}) \neq (0, 0, 0) \tag{2.12}$$

Suppose that $\bar{w} = 0$. Then, by (2.7), we have $\bar{u} = 0$ because $h_i(x) > 0, i = 1, \cdots, N$ for all $x \in X$. Hence, it follows from (2.6), (2.9), (2.10) and (2.12) that

$$\sum_{i \in I(\bar{x})} \bar{v}_i \nabla g_i(\bar{x}) = 0, \tag{2.13}$$

$$\sum_{i \in I(\bar{x})} \bar{v}_i > 0, \bar{v}_i \geq 0, i \in I(\bar{x}), \tag{2.14}$$

where $I(\bar{x}) = \{i | g_i(\bar{x}) = 0, 1 \leq i \leq m\}$. By Slater condition, there exists a point x' such that

$$g_i(x') < 0, i = 1, \cdots, m. \tag{2.15}$$

Since $g_i(x), i = 1, \cdots, m$ are convex, it follows from (2.15) that

$$\nabla g_i(\bar{x})^T (x' - \bar{x}) \leq g_i(x') - g_i(\bar{x}) < 0, i \in I(\bar{x}) \tag{2.16}$$

Letting $d = x' - \bar{x}$, from (2.16) and (2.14), we have $\left(\sum_{i \in I(\bar{x})} \bar{v}_i \nabla g_i(\bar{x}) \right)^T d < 0$, which contradicts (2.13). Thus, we have $\bar{w} > 0$.

Denoting $\frac{\bar{u}}{\bar{w}}$ and $\frac{\bar{v}}{\bar{w}}$ by $\bar{u}$ and $\bar{v}$ again, respectively, we see that (2.7) is equivalent to (2.4) and so (2.8) is equivalent to (2.5) because $\bar{u}_i > 0, i = 1, \cdots, N$ by (2.4).

Given $u = \bar{u}$ and $\beta = \bar{\beta}$, (2.6), (2.9) and (2.10) is just the KKT condition for the problem (2.3). Since the problem (2.3) is convex programming for parameters $u > 0$ and $\beta \geq 0$, the KKT condition is also sufficient optimality condition and then $\bar{x}$ is the solution of (2.3) for $u = \bar{u}$ and $\beta = \bar{\beta}$. □

Remark 2.1. Consider the maximization problem

$$\max F(x) = \sum_{i=1}^{N} F_i(x),$$

$$\text{subject to} \quad g_i(x) \leq 0, i = 1, \cdots, m,$$
$$x \in R^n$$

where $F_i(x) = \frac{f_i(x)}{h_i(x)}, i = 1, \cdots, N$ and $f_i(x), -g_i(x)$ and $-h_i(x)$ are concave functions, and $f_i(x) \geq 0, h_i(x) > 0, i = 1, \cdots, N$ in the feasible set X

The above problem is equivalent to the following problem.

$$\max F(x) = \sum_{i=1}^{N} \beta_i,$$

$$\text{subject to} \quad F_i(x) \geq \beta_i, i = 1, ..., N,$$
$$g_i(x) \leq 0, i = 1, ..., m,$$
$$x \in R^n$$

Then we can obtain the same result as the Lemma 2.1: If $(\bar{x}, \bar{\beta})$ is the solution of the above maximization problem , then there exist $\bar{u}_i, i = 1, \cdots, N$ such that $\bar{x}$ is a solution of the following problem for $u = \bar{u}$ and $\beta = \bar{\beta}$.

$$\max \sum_{i=1}^{N} u_i(f_i(x) - \beta_i h_i(x)),$$

$$\text{subject to} \quad g_i(x) \leq 0, i = 1, ..., m,$$
$$x \in R^n$$

And $\bar{x}$ also satisfies the system of equations (2.4) and (2.5) for $u = \bar{u}$ and $\beta = \bar{\beta}$.

The above problem is convex programming for given parameters $\beta \geq 0$ and $u > 0$. In what follows, all the results for the minimization problem hold true for the maximization problem. □

Let $\alpha = (\beta, u)$ denote parameter vector and let

$$\Omega = \{\alpha = (\beta, u) \in R^{2N} | 0 \leq \beta \leq \beta^u, 0 < u^l \leq u \leq u^u\},$$

where

$$\beta^u = (\beta_1^u, \cdots, \beta_N^u), u^l = (u_1^l, \cdots, u_N^l), u^u = (u_1^u, \cdots, u_N^u), \beta_i^u = \max_{x \in X} \frac{f_i(x)}{h_i(x)}, u_i^u = \max_{x \in X} \frac{1}{h_i(x)},$$

and

$$u_i^l = \min_{x \in X} \frac{1}{h_i(x)}, i = 1, \cdots, N.$$

Remark 2.2. The β^u, u^l and u^u involved in the definition of Ω are needed only for theoretical consideration and they are not needed in solving the problem (2.1) at all. □

Let $x(\alpha)$ be the solution of the problem (2.3) for fixed $\alpha \in \Omega$ and let

$$\varphi(\alpha) = \sum_{i=1}^{N} u_i(f_i(x(\alpha)) - \beta_i h_i(x(\alpha))) \tag{2.17}$$

$$\psi_i^1(\alpha) = -f_i(x(\alpha)) + \beta_i h_i(x(\alpha)), i = 1, \cdots, N, \tag{2.18}$$

$$\psi_i^2(\alpha) = -1 + u_i h_i(x(\alpha)), i = 1, \cdots, N \tag{2.19}$$

Let

$$\hat{\Omega} = \{\alpha \in \Omega | \psi_i^1(\alpha) = 0, \psi_i^2(\alpha) = 0, i = 1, \cdots, N\}.$$

and

$$\psi_i^1(\alpha) = 0, i = 1, \cdots, N, \tag{2.20}$$

$$\psi_i^2(\alpha) = 0, i = 1, \cdots, N \tag{2.21}$$

Corollary 2.1. *If $\bar{x}$ is the solution for the problem (2.1), then there exists $\alpha \in \Omega$ such that $\bar{x} = x(\alpha)$ and $\alpha \in \hat{\Omega}$.*

Proof. The proof follows from Lemma 2.1 because (2.5) and (2.4) implies (2.20) and (2.21), respectively. □

Corollary 2.2. *If the problem (2.1) has a solution $\bar{x}$, then $\hat{\Omega} \neq \emptyset$. And If $\hat{\Omega}$ consists of only one point α^*, i.e. $\hat{\Omega} = \{\alpha^*\}$, then $x(\alpha^*)$ is the solution of the problem (2.1).*

Proof. It follows from Corollary 2.1 that $\hat{\Omega} \neq \emptyset$. If $\hat{\Omega} = \{\alpha^*\}$, the $x(\alpha^*)$ is the solution for the problem (2.1) because $\bar{x} = x(\alpha^*)$ by Corollary 2.1. □

It follows from Corollary 2.2 that the problem (2.1) is reduced to find the solution satisfying (2.20) and (2.21) among the $x(\alpha)$, the solutions of the problem (2.3) for parameter vector α. If $\hat{\Omega} = \{\alpha^*\}$, then solving the problem (2.1) is equivalent to finding the $\alpha^* \in \hat{\Omega}$.

In the next section, we shall establish the uniqueness of $\hat{\Omega}$ under some asumptions and propose an algorithm to find the unique element of $\hat{\Omega}$, i.e. to find the global optimal solution for the problem (2.1).

3 Algorithm and its convergence

Let

$$\psi_i(\alpha) = \psi_i^1(\alpha), \quad \psi_{N+i}(\alpha) = \psi_i^2(\alpha), i = 1, \cdots, N.$$

Then (2.20) and (2.21) can be rewritten as

$$\psi(\alpha) = 0 \tag{3.1}$$

Lemma 3.1. *If the problem (2.3) has a unique solution for each fixed $\alpha \in \Omega$ and the corresponding Lagrange multiplier vector is also unique, then function $\varphi(\alpha)$ defined by (2.17) are differentiable with α and we have*

$$\frac{\partial \psi_i^1(\alpha)}{\partial \beta_j} = \begin{cases} h_i(x(\alpha)) & \text{if } i = j \ , \\ 0 & \text{if } i \neq j. \end{cases} \tag{3.2}$$

$$\frac{\partial \psi_i^2(\alpha)}{\partial u_j} = \begin{cases} h_i(x(\alpha)) & \text{if } i = j \ , \\ 0 & \text{if } i \neq j. \end{cases} \tag{3.3}$$

Proof. The Lagrange function of the problem (2.3) is

$$l(x, \lambda, \alpha) = \sum_{i=1}^{N} u_i(f_i(x) - \beta_i h_i(x)) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

By the assumption on the problem (2.1), the problem (2.3) is convex programming and satisfies the Slater's regularity condition. Therefore, the optimal value function $\varphi(\alpha)$ of the problem (2.3) is differentiable with α and $\nabla\varphi(\alpha) = \nabla_\alpha l(x, \lambda, \alpha)$ by the corollary of Theorem 4 of [21], where

$$\nabla_\alpha l(x, \lambda, \alpha) = \left(\frac{\partial l}{\partial \beta_1}, \cdots, \frac{\partial l}{\partial \beta_N}, \frac{\partial l}{\partial u_1}, \cdots, \frac{\partial l}{\partial u_N} \right)^T,$$

$$\frac{\partial l}{\partial \beta_i} = -u_i h_i(x(\alpha)), \quad \frac{\partial l}{\partial u_i} = f_i(x(\alpha)) - \beta_i h_i(x(\alpha)), 1 = 1, \cdots, N.$$

In this case, it is easy to see that the (3.2) and (3.3) hold for the functions (2.18) and (2.19). $\square$

Lemma 3.2. *If $\psi(\alpha)$ is differentiable, it is strongly monotone with constant $\delta > 0$ in Ω, where*

$$\delta = \min_i \delta_i, \delta_i = \min_{x \in X} h_i(x), i = 1, \cdots, N$$

Proof. By (3.2) and (3.3), the Jacobian matrix of $\psi(\alpha)$ is as follows.

$$\psi'(\alpha) = \begin{pmatrix} h_1(x(\alpha)) & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & h_2(x(\alpha)) & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & h_N(x(\alpha)) & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & h_1(x(\alpha)) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & h_2(x(\alpha)) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & h_N(x(\alpha)) \end{pmatrix} \tag{3.4}$$

Since $x(\alpha) \in X$ and $h_i(x(\alpha)) > 0, i = 1, \cdots, N$, $\psi'(\alpha)$ is positive definite. Therefore, for any $d \in R^{2N}$, we have

$$d^T\psi'(\alpha)d = \sum_{i=1}^{N} d_i^2 h_i(x(\alpha)) + \sum_{i=1}^{N} d_{i+N}^2 h_i(x(\alpha)) =$$

$$= \sum_{i=1}^{N} (d_i^2 + d_{i+N}^2) h_i(x(\alpha)) \geq \sum_{i=1}^{N} (d_i^2 + d_{i+N}^2)\delta_i$$

$$\geq \delta \sum_{i=1}^{2N} d_i^2 = \delta \|d\|^2,$$

which completes the proof. □

Let

$$A_i(\alpha) = \frac{f_i(x(\alpha))}{h_i(x(\alpha))}, \quad A_{N+i}(\alpha) = \frac{1}{h_i(x(\alpha))}, i = 1, \cdots, N$$

Lemma 3.3. *The equation $\psi(\alpha) = 0$ is equivalent to the equation $\alpha = A(\alpha)$. If the problem (2.1) has an optimal solution, the equation (3.1) has at least one solution in Ω.*

Proof. The first proposition is obvious from the definition of $A(\alpha)$ and $\psi(\alpha)$. It follows from Corollary 2.1 that there is $\alpha \in \Omega$ such that $\psi(\alpha) = 0$. □

Let's introduce a mapping $B(\alpha) = \pi_\Omega(\alpha - \lambda\psi(\alpha))$ to establish the existence and uniqueness of the solution of (3.1), where λ is a positive constant and $\pi_\Omega(a)$ denotes the projection of a onto Ω, i.e. $\pi_\Omega(a)$ is the solution of the following problem:

$$\min \|x - a\|^2, x \in \Omega,$$

Lemma 3.4. *For all $a \in R^n$ and for all $x \in \Omega$,*

$$(\pi_\Omega(a) - a)^T(x - \pi_\Omega(a)) \geq 0, \tag{3.5}$$

and for all $a, b \in R^n$

$$\|\pi_\Omega(a) - \pi_\Omega(b)\| \leq \|a - b\| \tag{3.6}$$

Proof. See Kinderlehrer and Stampcchia (1980). □

Lemma 3.5. *Assume that $\psi(\alpha)$ is differentiable and satisfies the Lipschitz condition with the constant M in Ω. Then $B : \Omega \to \Omega$ is a contractive mapping for all $\lambda \in (0, 2\delta/M^2)$.*

Proof. Since $\psi(\alpha)$ is strongly monotone with the constant δ by Lemma 3.2,

$$(\psi(\alpha') - \psi(\alpha))^T(\alpha' - \alpha) \geq \delta\|\alpha' - \alpha\|^2. \tag{3.7}$$

By (3.6), (3.7) and the Lipschitz condition, we have

$$\begin{gathered}
\|B(\alpha') - B(\alpha)\|^2 = \|\pi_\Omega(\alpha' - \lambda\psi(\alpha')) - \pi_\Omega(\alpha - \lambda\psi(\alpha))\|^2 \\
\leq \|\alpha' - \lambda\psi(\alpha') - (\alpha - \lambda\psi(\alpha))\|^2 = \|(\alpha' - \alpha) - \lambda(\psi(\alpha') - \psi(\alpha))\|^2 \\
= \|\alpha' - \alpha\|^2 - 2\lambda(\alpha' - \alpha)^T(\psi(\alpha') - \psi(\alpha)) + \lambda^2\|\psi(\alpha') - \psi(\alpha)\|^2 \\
\leq \|\alpha' - \alpha\|^2 - 2\lambda\delta\|\alpha' - \alpha\|^2 + \lambda^2\|\psi(\alpha') - \psi(\alpha)\|^2 \\
\leq (1 - 2\lambda\delta + \lambda^2 M^2)\|\alpha' - \alpha\|^2,
\end{gathered}$$

which implies that

$$\|B(\alpha') - B(\alpha)\| \leq q\|\alpha' - \alpha\|$$

for all $\lambda \in (0, 2\delta/M^2)$, where $q = \sqrt{1 - 2\lambda\delta + \lambda^2 M^2} < 1$. □

Theorem 3.1. *Assume that the problem (2.1) has an optimal solution, and that $\psi(\alpha)$ is differentiable and satisfies the Lipschitz condition in Ω. The equation $\alpha = B(\alpha)$ for all $\lambda \in (0, 2\delta/M^2)$ is equivalent to the equation $\psi(\alpha) = 0$ and the equation (3.1) has a unique solution.*

Proof. By Lemma 3.5 and the contractive mapping principle, $B(\alpha)$ has only one fixed point α^* in Ω for all $\lambda \in (0, 2\delta/M^2)$, i.e. $B(\alpha^*) = \alpha^*$. From (3.5), we have

$$\begin{gathered}
[\pi_\Omega(\alpha^* - \lambda\psi(\alpha^*)) - (\alpha^* - \lambda\psi(\alpha^*))]^T[\alpha - \pi_\Omega(\alpha^* - \lambda\psi(\alpha^*)) = \\
= [\alpha^* - (\alpha^* - \lambda\psi(\alpha^*))]^T(\alpha - \alpha^*) \geq 0
\end{gathered}$$

for $\alpha \in \Omega$, i.e. for every $\alpha \in \Omega$, we have

$$\psi(\alpha^*)^T(\alpha - \alpha^*) \geq 0. \tag{3.8}$$

Since the problem (2.1) has an optimal solution, there is a $\bar{\alpha} \in \Omega$ such that $\psi(\bar{\alpha}) = 0$ by Lemma 3.3. Then, it follows from (3.8) that

$$[\psi(\alpha^*) - \psi(\bar{\alpha})]^T(\bar{\alpha} - \alpha^*) \geq 0. \tag{3.9}$$

By Lemma 3.2, $\psi(\alpha)$ is strongly monotone with the constant δ and so we have

$$[\psi(\alpha^*) - \psi(\bar{\alpha})]^T(\alpha^* - \bar{\alpha}) \geq \delta\|\alpha^* - \bar{\alpha}\|^2.$$

This inequality and (3.9) together implies $0 \leq [\psi(\alpha^*) - \psi(\bar{\alpha})]^T(\bar{\alpha} - \alpha^*) \leq -\delta\|\alpha^* - \bar{\alpha}\|^2$. Therefore, $\|\alpha^* - \bar{\alpha}\| = 0$, i.e. $\alpha^* = \bar{\alpha}$. Thus , $\psi(\alpha^*) = 0$, which means that α^* is a solution of (3.1).

From the definition of $B(\alpha)$, it is obvious that $\alpha \in \Omega$ such that $\psi(\alpha) = 0$ is also the solution of $\alpha = B(\alpha)$.

Suppose that $\alpha' \in \Omega$ is a solution of (3.1) that is different from α^*. Then, $\alpha' = B(\alpha')$, which is a contradiction because α^* is a unique fixed point of $B(\alpha)$ in Ω. Thus, α^* is a unique solution of the equation (3.1). □

Corollary 3.1. *Assume that the problem (2.1) has an optimal solution, and that $\psi(\alpha)$ is differentiable and satisfies the Lipschitz condition in Ω. The equation $\alpha = A(\alpha)$ has a unique solution which is the unique solution of the equation (3.1).*

Proof. By Lemma 3.3, $\alpha = A(\alpha)$ is eqivalent to $\psi(\alpha) = 0$, which is equivalent to $\alpha = B(\alpha)$ by Theorem 3.1. Therefore, it follows from Theorem 3.1 that equation $\alpha = A(\alpha)$ has a unique solution which is the unique solution of the equation (3.1). □

As seen in Lemma 3.3, α is a fixed point of $A(\alpha)$ if and only if α is a root of $\psi(\alpha)$. We can construct the following simple iterative method to find a fixed point of $A(\alpha)$.

$$\beta^{k+1} = \left(\frac{f_1(x^k)}{h_1(x^k)}, \cdots, \frac{f_N(x^k)}{h_N(x^k)}\right)^T, u^{k+1} = \left(\frac{1}{h_1(x^k)}, \cdots, \frac{1}{h_N(x^k)}\right)^T, \tag{3.10}$$

where $x^k = x(\alpha^k)$.

Corollary 3.2. *The algoritm (3.10) is just the Newton method to solve the equation (3.1) and its local rate of convergence is superlinear or quadratic.*

Proof. The Newton method for the equation (3.1) is as following.

$$\alpha^{k+1} = \alpha^k - [\psi'(\alpha^k)]^{-1}\psi(\alpha^k) \tag{3.11}$$

By (3.4), (2.18) and (2.19), the right-hand side of (3.11) is equal to $A(\alpha^k)$, i.e. the right-hand side of (3.10). Therefore, (3.11) means $\alpha^{k+1} = A(\alpha^k)$, that is, the simple iterative method to find a fixed point of $A(\alpha)$ is just the Newton method for solving the equation (3.1). Hence, the algorithm (3.10) has local superlinear or quadratic convergence rate. □

Theorem 3.2. *Assume that the problem (2.1) has an optimal solution, and that $\psi(\alpha)$ is differentiable and satisfies the Lipschitz condition in Ω. And suppose that there exists a $L > 0$ such that for every $\alpha, \alpha' \in \Omega$*

$$\|\psi'(\alpha) - \psi'(\alpha')\| \le L\|\alpha - \alpha'\|, \tag{3.12}$$

and that there exists a $\tilde{M} > 0$ such that for every $\alpha \in \Omega$

$$\|[\psi'(\alpha)]^{-1}\| \le \tilde{M}. \tag{3.13}$$

Let

$$\alpha^{k+1} = \alpha^k + \lambda_k p^k, \qquad p^k = -[\psi'(\alpha^k)]^{-1}\psi(\alpha^k), \tag{3.14}$$

where λ_k is the greatest ξ^i satisfying

$$\|\psi(\alpha^k + \xi^i p^k)\| \leq (1 - \varepsilon\xi^i)\|\psi(\alpha^k)\| \tag{3.15}$$

and $i \in \{0, 1, 2, \cdots\}$, $\quad \xi \in (0,1)$, $\quad \varepsilon \in (0,1)$. Then, the modified Newton method defined by (3.14) and (3.15) converges to the unique solution α^* of $\psi(\alpha)$ with linear rate for any starting point $\alpha^0 \in \Omega$ and the rate in the neighborhood of the solution is quadratic.

Proof. We have already shown the existence and uniqueness of the solution to the equation $\psi(\alpha) = 0, \quad \alpha \in \Omega$, above. If there is k such that $\psi(\alpha^k) = 0$, then α^k is a solution. So, it is assumed that $\psi(\alpha^k) \neq 0$ for every k . For $\lambda \in [0,1]$, we have the following by the Newton-Leibnitz formula and (3.12).

$$\begin{aligned} \|\psi(\alpha^k + \lambda p^k)\| &= \left\| \psi(\alpha^k) + \lambda \int_0^1 \psi'(\alpha^k + \theta\lambda p^k) p^k d\theta \right\| \\ &= \left\| \psi(\alpha^k) + \lambda \int_0^1 \left[\psi'(\alpha^k + \theta\lambda p^k) - \psi'(\alpha^k)\right] p^k d\theta - \lambda\psi(\alpha^k) \right\| \\ &\leq (1-\lambda)\|\psi(\alpha^k)\| + \lambda^2 L \|p^k\|^2, \end{aligned} \tag{3.16}$$

where we took account of relation

$$\|\psi'(\alpha^k + \theta\lambda p^k) - \psi'(\alpha^k)\| \leq L\theta\lambda p^k$$

by the Lipschitz condition, and $\psi'(\alpha^k)p^k = -\psi(\alpha^k)$ by (3.14) . In view of (3.13), it follows from (3.14) and (3.16) that

$$\|\psi(\alpha^k + \lambda p^k)\| \leq \left[1 - \lambda(1 - \lambda L\tilde{M}^2)\|\psi(\alpha^k)\|\right] \|\psi(\alpha^k)\|.$$

Letting $\bar{\lambda}_k = \frac{1-\varepsilon}{L\tilde{M}^2\|\psi(\alpha^k)\|}$, we have

$$\|\psi(\alpha^k + \lambda p^k)\| \leq (1 - \varepsilon\lambda)\|\psi(\alpha^k)\| \tag{3.17}$$

for every $\lambda \in (0, \min\{1, \bar{\lambda_k}\})$. Then, (3.15), the definition of λ_k, implies that

$$\lambda_k \geq \min\{1, \xi\bar{\lambda_k}\} \tag{3.18}$$

Since it follows from (3.17) that

$$\|\psi(\alpha^{k+1})\| \leq (1 - \varepsilon\lambda_k)\|\psi(\alpha^k)\|, \tag{3.19}$$

$\{\|\psi(\alpha^k)\|\}$ is a monotonically decreasing sequence and so $\bar{\lambda_k}$ increases monotonically. From (3.18), it follows that

$$1 - \varepsilon\lambda_k \leq 1 - \varepsilon\min\{1, \xi\bar{\lambda_k}\} \leq 1 - \varepsilon\min\{1, \xi\bar{\lambda_0}\}.$$

Letting $q = 1 - \varepsilon \min\{1, \xi \bar{\lambda_0}\}$, then $q < 1$ and by (3.19) we have

$$\|\psi(\alpha^k)\| \leq q^k \|\psi(\alpha^0)\| \tag{3.20}$$

Therefore, $\{\|\psi(\alpha^k)\|\}$ converges to zero with linear rate, which means that converges to a solution of $\psi(\alpha) = 0$ for any starting point $\alpha^0 \in \Omega$. Since $\|\psi(\alpha^k)\| \to 0$ as $k \to \infty$, we have $\bar{\lambda_k} \to \infty$ as $k \to \infty$ and so there exists such k_0 that $\lambda_k = 1$ for every $k \geq k_0$. Thus, in this case, (3.14) becomes the Newton method and the rate of convergence in the neighborhood of the solution is quadratic by the Lipschitz property (3.12). □

Let us consider sufficient condition under which the assumptions of Theorem 3.2 are satisfied.

Theorem 3.3. *If $h_i(x), i = 1, \cdots, N$ and $x(\alpha)$ are Lipschitz continuous in the feasible set X and Ω, respectively, then $\psi'(\alpha)$ is Lipschitz continuous in Ω. And $[\psi'(\alpha)]^{-1}$ is bounded.*

Proof. By the Lipschitz continuity of $h_i(x)$ and $x(\alpha)$, there is a $L_i > 0$ and $C > 0$ such that, for all $\alpha', \alpha \in \Omega$,

$$|h_i(x(\alpha)) - h_i(x(\alpha'))| \leq L_i \|x(\alpha) - x(\alpha')\|, \tag{3.21}$$

$$\|x(\alpha) - x(\alpha')\| \leq C\|\alpha - \alpha'\| \tag{3.22}$$

Since, by (3.2) and (3.3),

$$\psi'(\alpha) - \psi'(\alpha') = \begin{bmatrix} diag\,(h_i(x(\alpha)) - h_i(x(\alpha'))) & 0 \\ 0 & diag\,(h_i(x(\alpha)) - h_i(x(\alpha'))) \end{bmatrix},$$

there exists a constant $L > 0$ by (3.21) and (3.22) such that $\|\psi'(\alpha) - \psi'(\alpha')\| \leq L\|\alpha - \alpha'\|$, which means the the Lipschitz continuity of $\psi'(\alpha)$.

It is easy to see

$$[\psi'(\alpha)]^{-1} = \begin{bmatrix} diag(\frac{1}{h_i(x(\alpha))}) & 0 \\ 0 & diag(\frac{1}{h_i(x(\alpha))}) \end{bmatrix}.$$

Thus, from $\frac{1}{h_i(x(\alpha))}) \leq u_i^u, \quad i = 1, \cdots, N$, it follows that there exists a $\tilde{M} > 0$ such that for every $\alpha \in \Omega$

$$\|[\psi'(\alpha)]^{-1}\| \leq \tilde{M},$$

that is, (3.13) is satisfied. □

The modified Newton method defined by

$$\alpha^{k+1} = \alpha^k - \lambda_k[\psi'(\alpha^k)]^{-1}\psi(\alpha^k)]$$

can be rewritten component-wise as following.

$$\beta_i^{k+1} = \beta_i^k - \frac{\lambda_k}{h_i(x^k)}[\beta_i^k h_i(x^k) - f_i(x^k)] = (1 - \lambda_k)\beta_i^k + \lambda_k \frac{f_i(x^k)}{h_i(x^k)}, \quad i = 1, \cdots, N \tag{3.23}$$

$$u_i^{k+1} = u_i^k - \frac{\lambda_k}{h_i(x^k)}[u_i^k h_i(x^k) - 1] = (1-\lambda_k)u_i^k + \lambda_k \frac{1}{h_i(x^k)}, \quad i = 1, \cdots, N \tag{3.24}$$

On the basis of the above consideration, we construct an algorithm to find global solution of the problem (2.1) as following.

[Algorithm MN]

Step 0. Choose $\xi \in (0,1), \quad \varepsilon \in (0,1)$ and $y^0 \in X$. Let

$$\beta_i^0 = \frac{f_i(y^0)}{h_i(y^0)}, \quad u_i^0 = \frac{1}{h_i(y^0)}, \quad i = 1, \cdots, N, \quad k = 0,$$

$$\beta^k = (\beta_1^k, \cdots, \beta_N^k), \quad u^k = (u_1^k, \cdots, u_N^k).$$

Step 1. Find a solution $x^k = x(\alpha^k)$ of the problem

$$min \sum_{i=1}^N u_i^k \left(f_i(x) - \beta_i^k h_i(x)\right),$$

$$\text{subject to} \quad x \in X.$$

for $\alpha^k = (\beta^k, u^k)$.

Step 2. If $\psi(\alpha^k) = 0$, then x^k is a global solution and so stop the algorithm. Otherwise, let i_k denote the smallest integer among $i \in \{0, 1, 2, \cdots\}$ satisfying

$$\|\psi(\alpha^k + \xi^i p^k)\| \leq (1 - \varepsilon\xi^i)\|\psi(\alpha^k)\|$$

and let

$$\lambda_k = \xi^{i_k}, \quad \alpha^{k+1} = \alpha^k + \lambda_k p^k, \quad p^k = -[\psi'(\alpha^k)]^{-1}\psi(\alpha^k)$$

Step 3. Let $k = k + 1$ and go to step 1. $\square$

Remark 3.1. For the sum-of-ratios maximization problem, the following maximization problem should be solved for $\alpha = (\beta, u)$ in the step 1 and the step 2 of Algorithm MN.

$$\begin{aligned} \max \sum_{i=1}^N u_i(f_i(x) - \beta_i h_i(x)), \\ \text{subject to} \quad x \in X. \qquad \square \end{aligned} \tag{3.25}$$

The step 2 of the algorithm MN is just the modified Newton method (3.14), (3.15). Therefore, the detailed update formula for $\alpha^k = (\beta^k, u^k)$ are (3.23) and (3.24). If the stepsize $\lambda_k = \xi^{i_k}$ in the step 2 is replaced by $\lambda_k = 1$, the step 2 is just the Newton method (3.11) and so we denote the algorithm MN by algorithm N in this case. As shown in Theorem 3.2, the Algorithm MN has global linear and local quadratic rate of convergence.

4 Numerical experiments

Our algorithms have been implemented by 1.5GHZ 20GB 256MB PC in the MATLAB 7.8 environment using the optimization toolbox. We made use of $|\psi(\alpha^k)| < 1e{-}6$ as the stopping criterion in experiments below.

Problem 1.

$$\max \quad \frac{x_1}{x_1^2 + x_2^2 + 1} + \frac{x_2}{x_1 + x_2 + 1},$$
$$\text{subject to} \quad x_1 + x_2 \le 1, \quad x_1 \ge 0, \quad x_2 \ge 0$$

The optimal value of the problem 1 is 0.5958 . Our experiments had been carried out by both the algorithm N and algorithm MN.

The performance of the algorithm N and MN for the problem 1 is given in the following table, where " titer" is the total amount of the convex programming problem solved during the algorithm's work and "iter" is the amount of iterations of the algorithm.

Initial point	N	MN	
	iter	titer	iter
(0,0)	7	8	7
random	7	9	7

In all experiments, we obtained the global solution with the accuracy of four-digits after 5 iterations. The algorithm MN required to solve twice the problem (3.25) in the first iteration and its stepsize had been always equal to 1 since the second iteration, that is, the algorithm MN turned to the algorithm N after the second iteration.

Problem 2.

$$\max \quad \frac{x_1}{x_1^2 + 1} + \frac{x_2}{x_2 + 1},$$
$$\text{subject to} \quad x_1 + x_2 \le 1, \quad x_1 \ge 0, \quad x_2 \ge 0$$

The optimal solution and optimal value are $x^* = (\frac{1}{2}, \frac{1}{2})$ and $\frac{4}{5}$, respectively.

Initial point	N	MN	
	iter	titer	iter
(0,0)	2	14	8
random	$\times$	5	3

In the above table, "×" denotes that the algorithm N faild to find any solution for all of 100 runs.

Problem 3 [9].

$$\max \quad \frac{-x_1^2+3x_1-x_2^2+3x_2+3.5}{x_1+1}+\frac{x_2}{x_1^2-2x_1+x_2^2-8x_2+20},$$
$$\text{subject to} \quad 2x_1+x_2\le 6, \quad 3x_1+x_2\le 8, \quad x_1-x_2\le 1, x_1\ge 1, \quad x_2\ge 1.$$

The optimal solution and optimal value are (1, 1.743839) and 4.060819 , respectively. The result of numerical experiments for the problem 3 is shown in the following table.

Initial point	N	MN	
	iter	titer	iter
(1,1)	8	10	8
(0,0)	8	9	8
random	8	9	8

In [9], this problem was solved by branch-and-bound algorithm in which after 27 iterations of the algorithm, an ε-global solution (1.0200, 1.7100) with value 4.0319 was found for $\varepsilon = 0.05$. In this example, unlike in many other branch-and-bound applications, the feasible solution that was shown to be globally ε-optimal was not discovered until very late in the search process. In contrast with the result of [9], in all of 100 experiments with random initial point, the both of algorithm N and MN obtained the global solution with the accuracy of four-digits after 6 iterations. The algorithm MN required to solve twice or thirdly the problem (3.25) in the first iteration and its stepsize had been always equal to 1 since the second iteration, that is, the algorithm MN turned to the algorithm N after the second iteration.

This problem demonstrates the advantage of our algorithm compared with the previous algorithms based on the branch-and-bound method.

Problem 4 [14].

$$\max \quad \frac{3x_1+5x_2+3x_3+50}{3x_1+4x_2+5x_3+50}+\frac{3x_1+4x_2+50}{4x_1+3x_2+2x_3+50}+\frac{4x_1+2x_2+4x_3+50}{5x_1+4x_2+3x_3+50},$$
$$\text{subject to} \quad 6x_1+3x_2+3x_3\le 10, \quad 10x_1+3x_2+8x_3\le 10, \quad x_1,x_2,x_3\ge 0.$$

The optimal solution and optimal value are (0, 3.3333, 0) and 3.002924 , respectively. In [14], after 21 iterations, a 10^{-3} - optimal solution to Problem 4, (0.000, 0.000, 1.250) was obtained. The amount of iteration in [13] was 25. The performance of our algorithms for the problem 4 is shown in the following table.

Initial	N	MN	
point	iter	titer	iter
(1,1,1)	2	3	2
(0,0,0)	2	4	2
random	2	3	2

Problem 5 [14].

$$\max \quad \frac{4x_1+3x_2+3x_3+50}{3x_2+3x_3+50}+\frac{3x_1+4x_3+50}{4x_1+4x_2+5x_3+50}+\frac{x_1+2x_2+5x_3+50}{x_1+5x_2+5x_3+50}+\frac{x_1+2x_2+4x_3+50}{5x_2+4x_3+50},$$

$$\text{subject to} \quad 2x_1+x_2+5x_3 \le 10, \quad x_1+6x_2+3x_3 \le 10,$$

$$5x_1+9x_2+2x_3 \le 10, \quad 9x_1+7x_2+3x_3 \le 10, \quad x_1,x_2,x_3 \ge 0.$$

The optimal solution and optimal value are (1.1111, 0, 0) and 4.090703 , respectively. In [14], 10^{-3} -optimal solution, (0.000, 1.111, 0.000), to Problem 5 in 31 iterations and 15 branching operations. The amount of iteration in [13] was 3. The result of numerical experiments for the problem 5 is shown in the following table.

Initial	N	MN	
point	iter	titer	iter
(1,1,1)	2	3	2
(0,0,0)	2	3	2
random	2	3	2

Problem 6 [12].

$$\min \quad \frac{x_1^2+x_2^2+2x_1x_3}{x_3^2+5x_1x_2}+\frac{x_1+1}{x_1^2-2x_1+x_2^2-8x_2+20},$$

$$\text{subject to} \quad x_1^2+x_2^2+x_3 \le 5, \quad (x_1-2)^2+x_2^2+x_3^2 \le 5,$$

$$1 \le x_1 \le 3, \quad 1 \le x_2 \le 3, \quad 1 \le x_3 \le 2.$$

The optimal solution and optimal value are (1, 1, 1) and 0.8333333 , respectively. The solution and value obtained in [12] are (1.32547, 1.42900, 1.20109) and 0.9020, respectively, and the amount of iteration is 84. The optimal solution was found after 10 iterations in [10]. The numerical experiment results of our algorithms for the problem 6 is shown in the following table.

Initial	N	MN	
point	iter	titer	iter
(0,0,0)	2	4	2
random	2	4	2

Problem 7 [10].

$$\min \quad \frac{-x_1^2+3x_1-x_2^2+3x_2+3.5}{x_1+1}+\frac{x_2}{x_1^2+2x_1+x_2^2-8x_2+20},$$

$$\text{subject to} \quad 2x_1+x_2 \leq 6, \quad 3x_1+x_2 \leq 8, \quad x_1-x_2 \leq 1$$

$$0.1 \leq x_1, x_2 \leq 3$$

The optimal solution and optimal value are (0.1, 0.1) and 3.7142402 , respectively. The optimal solution was found after 19 iterations in [10]. The numerical experiment results of our algorithms for the problem 7 is shown in the following table.

Initial	N	MN	
point	iter	titer	iter
(0,0)	2	5	2
(1,1)	2	4	2
random	2	4	2

Problem 8 [8].

$$\min \quad \frac{x_1^2-4x_1+2x_2^2-8x_2+3x_3^2-12x_3-56}{x_1^2-2x_1+x_2^2-2x_2+x_3+20}+\frac{2x_1^2-16x_1+x_2^2-8x_2-2}{2x_1+4x_2+6x_3},$$

$$\text{subject to} \quad x_1+x_2+x_3 \leq 10, \quad -x_1-x_2+x_3 \leq 4,$$

$$1 \leq x_1, x_2, x_3$$

The optimal solution and optimal value are (1.82160941, 1, 1) and -6.11983418 , respectively. With ε=0.01, a global ε -optimal solution (1.81, 1.00, 1.00) with ε-optimal value -6.12 was found after 24 iterations in [8]. It took time of 28.4 seconds for finding the optimal value in [11]. The numerical experiment results of our algorithms for the problem 8 is shown in the following table.

Initial point	N	MN	
	iter	titer	iter
(1,1,1)	4	6	4
random	4	6	4

Remark . Although the problem 6 and 7 do not satisfy the assumption on the problem (2.1) that numerators are convex and denominators are concave in all ratios, our Algorithm N and Algorithm MN found optimal solutions of the problems. This fact shows that our algorithm can find global solution for more general nonlinear sum of ratios problems. □

The following table shows optimal solutions, optimal values and computational times of Algorithm MN for Problem 1 $\sim$ Problem 8.

Example number	optimal solution			optimal value	time
	x_1	x_2	x_3	F^*	(second)
1	0.6390292	0.3609708		0.5958011	0.0586
2	0.5000524	0.4999476		0.7999999	0.0975
3	1.0000000	1.7438387	1.0000000	4.0608191	0.0992
4	0.0000000	3.3333333	0.0000000	3.0029239	0.0563
5	1.1111111	0.0000000	0.0000000	4.0907029	0.0594
6	1.0000000	1.0000000	1.0000000	0.8333333	0.1015
7	0.1000000	0.1000000		3.7142402	0.0898
8	1.8216094	1.0000000	1.0000000	-6.1198342	0.0517

Problem 9. We carried out numerical experiments for randomly generated minimization problems as follows. The numerator and denominator of each term in objective function are

$$f_i(x) = \frac{1}{2}x^T A_{0i}x + q_{0i}^T x, \quad i = 1, ..., N$$

and

$$h_i(x) = c_i^T x, \quad i = 1, ..., N,$$

respectively. The feasible set is given by

$$X = \{x \in R^n | Ax \le b, \quad 0 \le x_i \le 5, \quad i = 1, ..., n\}$$

where $A_{0i} = U_i D_{0i} U_i^T, \quad U_i = Q_1 Q_2 Q_3, \quad i = 1, ..., N,$

$$Q_j = I - 2\frac{w_j w_j^T}{\|w_j\|^2}, \quad j = 1, 2, 3$$

and

$$w_1 = -i + \text{rand}(n, 1), \quad w_2 = -2i + \text{rand}(n, 1), \quad w_3 = -3i + \text{rand}(n, 1),$$

$$c_i = i - i \cdot \text{rand}(n, 1), \quad q_{0i} = i + i \cdot \text{rand}(n, 1),$$

$$A = -1 + 2 \cdot \text{rand}(5, n), \quad b = 2 + 3 \cdot \text{rand}(5, 1).$$

Starting with randomly generated point in $[1, 5]^n$, we carried out 200 runs of the algorithm N and MN for fixed n and N. Both of the algorithm N and MN were successful for all generated problems. The average performance is shown in the table below.

(i) algorithm N, i.e.(3.10)

n	N=5		N=10		N=50		N=100		N=200	
(dim)	iter	time(s)	iter	time(s)	iter	time(s)	iter	time(s)	iter	time(s)
10	6	0.2731	6	0.2753	6	0.2776	5	0.2844	5	0.2971
50	5	1.3288	5	1.3456	5	1.3764	5	1.7970	5	1.9313
100	5	4.4481	5	4.6139	5	4.6573	5	5.3741	5	5.6703
200	5	27.8647	5	27.9325	5	28.2704	6	29.1223	6	29.2663

(ii) algorithm MN, i.e.(3.14),(3.15)

n	N=5		N=10		N=50		N=100		N=200		
(dim)	iter	titer	iter	titer	iter	titer	iter	titer	iter	titer	time(s)
10	5	6	5	6	5	6	4	5	4	5	0.2984
50	5	6	4	5	4	5	4	5	4	5	1.1442
100	5	6	4	5	4	5	4	5	4	5	5.3757
200	5	6	4	5	5	6	4	5	4	5	29.4302

As shown in the above table, the amount of iterations of the algorithm N and the algorithm MN was independent of the amount of variables and fractional functions, and the run-time was proportional to the amount of variables. The experiments demonstrated that the algorithm N was successful for all problems but few problems , while the algorithm MN always gave global solution with any starting point.

5 Discussions

In this paper, we have presented a global optimization algorithm for the sum-of-ratios problem, which is non-convex fractional programming problem arising in computer graphics and finite element method. Our approach is based on transforming the sum-of-ratios problem into parametric convex programming problem and applying Newton-like algorithm to update parameters. The algorithm has global linear and local superlinear/quadratic rate of convergence. We have demonstrated that proposed algorithm always finds the global optimum with a fast convergence rate in practice by numerical experiments. In fact, the experiments showed that the optimal point is usually obtained within the first few iterations of the algorithm. So, we expect that our approach will solve successfully the several problems in multiview geometry formulated in [4].

References

[1] Bazaraa, M. S., Sherali, H. D., Shetty, C. M., *Nonlinear Programming: Theory and Algorithms,* John Wiley and Sons, New York, (2006)

[2] Carvalho, P. C. P., Figueiredo, L. H., Gomes, J. and Velho, L., *Mathematical Optimization in Computer Graphics and Vision,* Elsevier, Morgan Kaufmann Publishers (2008)

[3] Munson, T., *Mesh shape-quality optimization using the inverse mean-ratio metric,* Math. Program., 110(2007), 561-590

[4] Kahl, F., Agarwal1, S., Chandraker, M. K., Kriegman, D. and Belongie, S., *Practical Global Optimization for Multiview Geometry,* Int.J.Comput.Vis., 79(2008), 271-284

[5] Freund, R.W. and Jarre, F., *Solving the sum-of-ratios problem by an interior-point method,* J. Global Optimization, 19(2001), 83-102

[6] Tawarmalani, M. and Sahinidis, N.V., *Semidefinite relaxations of fractional programs via novel convexification techniques,* J. Global Optimization, 20(2001), 137-158

[7] Benson, H.P., *Solving Sum of Ratios Fractional Programs via Concave Minimization,* J.Optimization Theory and Application, 135(2007), 1-17

[8] Benson, H.P., *Using concave envelopes to globally solve the nonlinear sum-of-ratios problem,* J. Global Optimization, 22(2002), 343-364

[9] Benson H.P., *Global Optimization Algorithm for the Nonlinear Sum of Ratios Problem,* J.Optimization Theory and Application, 112(2002), 1-29

[10] Wang C.-F., Liu S.-Y. and Shen P.-P., *Global optimization for sum of geometric fractional functions,* Applied Mathematics and Computation 216 (2010), 2263-2270

[11] Wu W.-Y., Sheu R.-L. and Birbil S. I., *Solving the sum-of-ratios problem by a stochastic search algorithm,* J. Global Optimization, 42 (2008),91-109

[12] Qu S.J., Zhang K.C. and Zhao J.K., *An efficient algorithm for globally minimizing sum of quadratic ratios problem with nonconvex quadratic constraints,* Applied Mathematics and Computation 189 (2007), 1624-1636

[13] Shen P.-P. and Wang C.-F., *Global optimization for sum of linear ratios problem with coefficients,* Applied Mathematics and Computation 176 (2006,) 219-229

[14] Kuno, T., *A branch-and-bound algorithm for maximizing the sum of several linear ratios,* J. Global Optimization, 22 (2002), 155-174

[15] Schaible, S. and Shi, J., *Fractional programming: the sum-of- ratios case,* Opt. Meth. Soft., 18(2003), 219-229

[16] Hartley, R. and Sturm, P. *Triangulation,*CVIU 68(1997), 146-157

[17] Stewenius, H., Schaffalitzky, F. and Nister, D. *How hard is three-view triangulation really?*In: Int. Conf. Computer Vision. (2005), 686-693

[18] Kahl, F. and Henrion, D. *Globally optimal estimates for geometric reconstruction problems,* In: Int. Conf. Computer Vision, Beijing, China, (2005), 975-985

[19] Ke, Q. and Kanade, T. *Robust L_1- norm factorization in the presence of outliers and missing data by alternative convex programming,* In: CVPR. (2005), 739-746

[20] Kinderlehrer, D. and Stampcchia, G. *An introduction to variational inequalities and their applications*. Italiana: Academic Press (1980)

[21] Rockaffellar, R.T. *Directional differentiability of the optimal value function in a nonlinear programming problem,* Mathematical Programming Study, 21(1984), 213-226